

What will happen to UNESCO?

The British Government seems about to follow the United States into resignation from UNESCO. That will not be the end of the world (or even of UNESCO).

THE next few weeks will decide whether the UN Educational, Scientific and Cultural Organization (UNESCO) will survive in its present shape into the 1990s. And the outlook is far from bright. There now seems very little chance that the United States will go back on its decision of a year ago to pull out six weeks from now. The State Department's demand that the UNESCO administration should mend its ways has been partly met, but not sufficiently to persuade a president temperamentally biased towards withdrawal to change his mind. And there seems to be a strong constituency, among UNESCO's users in the United States, that last year's decision should be unscrambled (see p.300).

UNESCO's problem of managing without a quarter of its budget has now been further complicated by the chance that the United Kingdom will also, in the next few weeks, give notice of withdrawal. Such a decision, on the cards since the stern letter to UNESCO from the British government earlier in the year, has been made more likely by the tightness of the budget on which the Foreign and Commonwealth Office will have to live in the next three years. Even UNESCO's supporters in Britain, those who have held that £5 million a year is a small price to pay for continued membership of an organization from which withdrawal is bound to give widespread offence, will think differently when they recognize that the UNESCO subscription can be paid only at the expense of other useful activities — those of the British Council (the United Kingdom's own UNESCO), of the overseas broadcasts of the British Broadcasting Corporation or of direct foreign aid (for example to Ethiopia). So it is likely that the British government, anxious to clear its collective desk before the end of the year, will write a letter giving notice of withdrawal, put it in the mail and hope that the fuss will be dissipated during the coming holidays.

But what fuss? The obvious risks of withdrawal from UNESCO are political. To pull out of one UN organization is to put membership of all others in doubt, inviting the complaint that loyalty to international causes has been undermined. Because UNESCO has become widely regarded by the governments of developing countries as one of the chief vehicles for the transfer of the skills of advanced societies to those that lack them, those that withdraw will run the further risk of being thought indifferent to the cause of development. And because the row within UNESCO has been closely linked with personal criticism of the present director general, Mr Amadou-Mahtar M'Bow, enthusiastically the choice of the non-aligned countries on his appointment ten years ago, pulling out is bound to seem a slap in the face to those who voted for him in the first place.

Justification

The United States government is better able to meet these complaints head on than is the British; it can always point to the scale of its financial contributions to international organizations

of all kinds, the International Monetary Fund and the World Bank as well as the United Nations system. The British government has less to boast about, and is politically involved with many of UNESCO's fondest friends through the machinery of the British Commonwealth. If Britain does in the end withdraw, it had better pay close attention to the way in which its decision is justified.

Many of the complaints against UNESCO in the past several years are beside the point. Both the US and British governments say that UNESCO, with its remit to consider matters educational, scientific and cultural, should not be caught up in the political arguments that have dominated its general assemblies in recent years. But that is to ask too much. An organization whose terms of reference give it a proper interest in the care of the cultural heritage (and which has done valuable work in safeguarding the monuments along the Nile) cannot, for example, avoid also giving Greece an opportunity for attacking the British government because the Elgin marbles, which once adorned the Parthenon, are still on show in the British Museum. Similarly, it has from the start been inevitable that UNESCO should have been a battleground for conflicting interests over issues such as the control of copyright on publications (and the 1971 Paris amendment to the Berne Convention on copyright is both a feather in UNESCO's cap and a considerable victory for developing countries within UNESCO).

Given UNESCO's membership and constitution, such issues have been unavoidable from the start. Even Mr M'Bow's "new information order", a set of proposals whose immediate effects on the international news gathering agencies would have been disastrous, and which would also have legitimized the ambitions of governments to control what is said about them, was a proper question for UNESCO to consider. The complaint against UNESCO, and its management in particular, is that it seems not to have wished to lift a finger to avoid irreconcilable conflicts on these issues. Indeed, it has often seemed to relish them.

Blame

For this state of affairs, the developing countries must shoulder a substantial part of the blame. UNESCO is not by its constitution a supranational authority whose decisions are binding on its members, but rather a forum in which countries with something that others want may be persuaded to offer it. This is what happened with the revision of copyright protection, an international convention in whose negotiation UNESCO acted as a midwife, by which governments of developing countries were given limited rights to use educational materials if the original publishers are unable or unwilling to make them available. The difficulty that has arisen in the past few years is that the principle of one government, one vote, has gone to the heads of a majority of the membership. Arguments about issues such as the new information order are one thing (and should be tolerable), but proposals to spend a substantial part of communal funds supplied chiefly by those opposed to the proposal are another. One of the changes at UNESCO for which the United States has been pushing in the past year is the replacement of the present voting procedure by one in which those who pay the most have the most influence, but that would have been a cosmetic device for burying conflict, not for creating the atmosphere of consensus without

Science in the Soviet Union

THE special issue of *Nature* surveying science in the Soviet Union has been delayed until next month. □



which any communal fund must be misspent.

But why has the management allowed things to come to such a pass? That is a question that may always be fairly asked when a voluntary organization is in danger of being torn apart by internal conflicts. The usual answer is that the management has failed to spot trouble coming, and to head it off. That complaint unquestionably applies to UNESCO's management in the past few years. Mr M'Bow has been most often criticized, during UNESCO's troubled two years, for his tolerance of inefficiency, for his dictatorial ways with UNESCO's staff, for nepotism and because of his undoubted vanity. But the failing that matters is that he has not been able, or has not chosen, to function as the diplomat an organization such as UNESCO needs. Whatever members depart from UNESCO in the next few weeks, there is the strongest possible reason why Mr M'Bow should also quit — or be asked to go. UNESCO will be a substantial organization even if Britain follows the United States (and is followed by other members) into the cold. Especially with a shrunken budget, there is the strongest reason why UNESCO should be managed well.

Inevitably, it goes against the grain for governments wedded to the doctrine that it is injudicious (to say the least) to attack people and not policies to tell Mr M'Bow outright that he is an encumbrance to the organization he purports to head. Even so, it is mystifying that Mr George Shultz, when giving notice a year ago of the United States' intentions to withdraw, should have gone out of his way to emphasize that Mr M'Bow's misfortune is that he happens to have led UNESCO from a state of inconspicuous mediocrity to one of conspicuous failure. He would have been better served if the member governments now bent on withdrawing had told him sooner, and to his face, what they were always eager to tell others privately. He may now fairly complain that he was set up. But he has been a failure nevertheless.

By now, the principles on which reform should be sought are plain. Since it began, UNESCO has been wrongly organized for what it is now hoping to accomplish, the spending of a relatively modest \$200 million a year on good works scattered about the surface of the Earth. In effect, it is a modestly-sized charitable foundation whose terms of reference are international, but it is organized not for spending money efficiently but as if it were the kernel of an evolving academy of all the talents. The result is that the members of its staff are never sure whether they have been appointed in their capacity as scholars or as people who can channel funds to others who will be productive. And there are too many of them, by any set of job-descriptions, perhaps by an order of magnitude. It is laughably outrageous that UNESCO should spend two-thirds of its income on itself. When the present alarm about withdrawal, intended or merely threatened, is over, the remaining members should set about putting things to rights. The French government should be well placed to take the initiative.

Determination

What needs to be done is simply told. Inefficiency apart, UNESCO differs from other organizations. Ideally, its role should be to sponsor international intellectual projects that others have not attempted, being prepared to see many of these fail but resolute in its determination that those that succeed should be launched independently, with other people's funds, once their success is apparent. Instead, UNESCO has more often waited for other people's band-wagons to pass, seeking a small share in their sponsorship but a larger share of the credit. This, for example has been UNESCO's stance during the past quarter of a century in respect of the improvement of science teaching in secondary schools, where great opportunities have been missed for translating internationally programmes that have functioned well in member states. By contrast, UNESCO has embraced with too much enthusiasm schemes for teaching what is loosely called "peace" to unsuspecting teenagers throughout the world in circumstances when the most sophisticated educational systems are unable to provide models on which to practise what is being preached.

The conventional assumption is that none of these criticisms applies to UNESCO's scientific work, but even that claim is

shaky. It is, however, true that UNESCO is at present the conduit for continuing support for some important institutions, the International Council of Scientific Unions (ICSU) and the International Centre for Theoretical Physics (at Trieste), originally supported by the International Atomic Energy Agency. There is the strongest case why such institutions should not collapse, and why members thinking of withdrawing from UNESCO should ensure that their contributions to these central organizations should continue. UNESCO has also become the umbrella for a number of important coordinating research committees, the best known and most useful in oceanography. Thus the organization provides house-room and a secretariat for the International Oceanographic Commission, and also gathers oceanographic data collected by merchant vessels on routine voyages — tasks not at present done by other means. But UNESCO happens to be responsible only by accident, because it was the first to take on these tasks. As things have turned out, it would be more appropriate that such jobs should be done by ICSU or one of its member councils — and that the practical task of data-storage should be turned over to suitable laboratories working to contracts.

What might have been UNESCO's pride and joy in the past few years, however, the Man and Biosphere Programme (which is really an umbrella for a host of coordinated ecological projects scattered over the surface of the Earth, separately financed by national governments), is one to which the management in Paris has devoted much less enthusiasm than the case demands. But there is enough merit in UNESCO's scientific programme for the price of withdrawal from the organization to be an undertaking that the deserving parts of the programme should not be harmed.

Earlier this year, a committee of the US National Academy of Sciences under Walter A. Rosenblith of Massachusetts Institute of Technology made just this point to the State Department. Will the British government be prepared to follow suit? Its first reaction will be to say that if there is a chance of saving £5 million a year, it would prefer to keep the lot for other deserving causes. But this would be unacceptable, at least if worthwhile international institutions are to be endangered. Another unavoidable question is that of whether ICSU will be able to shoulder the extra burdens likely to be thrust on it even if the United States is the only member to pull out of UNESCO. The answer is that, at present, ICSU as such is not equipped to do much more than organize the committees necessary to ensure its continuing existence, but that situation need not persist. Indeed, there is a strong and quite separate case why, in the interests of the international community, ICSU should now be strengthened, and equipped with the resources necessary for it to play a stronger executive role in international science. But there is an obvious danger in supposing that ICSU might become a partial substitute for UNESCO — that ICSU might in the process itself be damaged, and that in any case there is no escaping some of the political difficulties that have made UNESCO uncomfortable in the past few years.

Reorganization

How, in these circumstances, should the British government decide? The question is not black and white. Pulling out regardless of the consequences would be foolish (and is unlikely). Pulling out without being prepared to pick up the pieces would similarly be wrong. Giving notice of withdrawal on the vague understanding that the decision might be reversed if UNESCO mends its ways would be fruitless, as the experience of the United States in the past year has shown. The best course is to couple a declaration of an intention to withdraw (not necessarily a year from now) with a prospectus describing how UNESCO might be organized efficiently, and how its reduced budget might nevertheless be spent effectively. The trouble, in the past year's bickering, is that it has concentrated on attempts to reform a well-intentioned organization absurdly ill-suited to its purposes by tinkering with the administration. What UNESCO needs is not to be told repeatedly why it is absurd, but to be helped to understand what it is for, and how it might function well. □

Fifth generation computers

Japan finds rivals treading uncomfortably close

Tokyo

JUDGING by the scale of Western governments' reactions to the fifth generation computer project, described by representatives to the project's international conference which ended in Tokyo last week, Japanese officials must be wondering if they would not have been better off keeping quiet about the whole subject.

Just three years ago, the project was launched in a storm of hyperbole that made it seem that Japan was about to seize world computer markets through a qualitative leap forward into a new world of super-intelligent computing systems. Now, Japan no longer looks the world leader as tough competing projects launched by the United Kingdom, the United States, France, Germany and the European Communities take off. Nor, despite earlier hopeful discussions between the United Kingdom and Japan, does there seem to be any longer a significant chance that cooperation will replace competition.

Strangely, the project that began it all, Professor Kazuhiro Fuchi's fifth generation computer project run from the Institute for New Generation Computing (ICOT), now looks academic and narrow, despite having been originally branded a "computing apocalypse" that would destroy Western nations' information technology industries. The explanation is Fuchi's continuing belief that the next phase of computer history will come from computers that use predicate logic to carry out inferences at high speed.

The language PROLOG, specially designed to handle such operations (see *Nature* 309, 198; 1984) has been taken as the project's starting point. With an inference mechanism at its core, the fifth generation project will try to use high speed mechanisms for retrieval from massive knowledge bases, permitting the computer to function as a highly intelligent "expert" system, while intelligent interfaces allow users to speak to their computers.

The intended results, very high speed inference and knowledge retrieval and intelligent man-machine interfaces, are what all the competing nations' projects aim at. But not everyone believes that predicate logic — and PROLOG — is the way to achieve them. As Daniel Bobrow of Xerox's Intelligent Systems Laboratory put it, "no single paradigm is appropriate to all problems. Just as there are many tools in a carpenter's tool box, there must be many in the programmer's kit." Fuchi sees many other approaches as being subsumed by logic programming (not to be confused with logical programming). If he

is right, his institute may well be way out in front in a couple of years; if not it could be left trying to "pry up nails with a screw-driver".

So far, anyway, ICOT has delivered on its promises. A working sequential inference machine was on display at the conference and will provide the essential "programming environment" (essentially equivalent to a laboratory bench and equipment) for further research. But from now on, the really hard problems begin — particularly that of parallel processing — where there are few clues as to which path or even which general philosophy will prove most profitable.

By contrast, Britain's Alvey Programme, although unashamedly inspired by ICOT's project, now looks quite catholic. It centres on four themes: software engineering, man-machine interfaces, intelligent knowledge-based systems and the development of very large scale integrated circuits, but admits of a variety of approaches. To counter the criticism best expressed by Robert Kowalski, one of PROLOG's inventors, that "it can be argued that it is harmful to support British research, because it is more likely to be exploited abroad than at home", all projects involve industrial participation and (almost) half industrial finance — even university-based projects have to find an industrial "uncle".

This in turn has led to an ICOT criticism of the Alvey project as "too incremental, too much aimed at short-term industrial application and lacking a basic research component" — exactly the complaints that have been flung at Japan's research for the past decade. And the industrial link has smashed hopes of cooperation between the United Kingdom and Japan. ICOT's project is 100 per cent supported by the government and ICOT will hold the patents.

With a budget of £350 million for the five years of the Alvey Programme, expenditure is running at double the rate of ICOT's project. But ICOT is only a tiny part of Japan's advanced computing effort: the telecommunications monopoly NTT's research laboratories have their own fifth generation project, and all of the large electronics manufacturers have artificial intelligence projects.

Only the United States, in total, is exceeding Japan's research efforts. Special projects launched there include the Defense Advanced Research Project Agency (DARPA)'s strategic computing programme (\$600 million over five years), the very high speed integrated circuit programme (\$100 million in 1984), other government programmes and new univer-

sity centres (\$100 million in 1984). There is also the private sector Microelectronics and Computer Technology Corporation and of course AT&T and IBM whose doings remain secret.

The West German response is more recent, stemming from the March 1984 government report on information technology which sets aside DM2,640 million over five years for information technology research. Collaboration between industry and the universities is again specially emphasized: key projects are on computer-aided design, new computer architectures and knowledge engineering.

France does not have a single grand plan, but has ambitions (see page 297). In 1983 the French initiated seven national projects, involving both industrial and public research laboratories, with themes broadly similar to those of the Alvey project. Now in preparation is a much longer-term set of fundamental research projects, centred on the Centre National de la Recherche Scientifique but also involving industry.

Finally, Esprit aims to overcome the European Communities' main weakness — fragmentation of research efforts — and rebuild the research base through "cooperation in precompetitive research and development". In 1984, 110 projects were being supported at 500 companies, institutes and universities, and expenditure of 200 million ECU approved.

All countries report that their biggest problem is finding trained staff — Alvey director Brian Oakley made in plain that anyone studying artificial intelligence will not be joining Britain's unemployed.

Alun Anderson

Apartheid telescoped

A PETITION signed by 156 UK astronomers has been handed to the Science and Engineering Research Council (SERC) calling for its links with the South African Astronomical Observatory (SAAO) to be cut, in the belief "that the continued funding (*sic*) of SAAO by SERC amounts to acquiescence in apartheid". SERC will not respond to the petition because, according to a spokeswoman, such decisions can be made by SERC only on the basis of scientific considerations.

SERC supports SAAO to the tune of about £250,000 a year together with travel and project grants to UK observers, who make considerable use of its facilities. Continuing use of the observatory is to be assessed next month as a matter of routine — though in the context of unusually tight financial constraints — during a "forward look" review by a sub-committee of SERC's Astronomy, Space and Radio Board. The controversy was last aired in 1982 when, as a result of a questionnaire circulated by SERC, it was reported that a majority of the community supported continued use of SAAO (see *Nature* 298, 698; 1982).

Philip Campbell

US research objectives

Keyworth's shopping list

Washington

MORE government support for research and training in biochemical engineering is urgently needed if the United States is to lead commercialization of the new biology, according to advice given by an expert panel to the White House Office of Science and Technology Policy (OSTP).

The panel, convened by the Committee on Science, Engineering and Public Policy (COSEPUP), says the United States lags in basic chemical engineering expertise and that if the position is not corrected, "the Japanese and the Europeans will certainly take the leading role".

The panel concludes that the government has a critical part to play in establishing US expertise in reactor design and purification methods and proposes a variety of grants and awards to be administered by federal agencies. The number of people graduating with higher degrees in biochemical engineering should be doubled or tripled from the present 60 per year.

The research briefing panel on chemical and process engineering for biotechnology is the most outspoken of 9 separate panels organized this year by COSEPUP. The briefings were given to OSTP officials, including director George Keyworth, during the autumn but published only last week.

COSEPUP briefings have over the past 3 years had a significant effect on OSTP policy; the proposed Engineering Research Centers, for example, stem directly from a COSEPUP initiative.

Dr Keyworth has this year taken a more active part in selecting subjects than previously. But while pleas for financial support for particular areas are taken seriously, according to officials, it is thought that some of the matters taken up by this year's briefing panels have already received special attention; in particular, the engineering research centres planned for some university engineering departments may go a long way towards meeting present needs.

A panel on computer architecture has urged immediate efforts to accelerate development of advanced software for the forthcoming generation of parallel processing computers: several indicators suggest that Japanese activities in this field exceed present US capability. And although the United States hopes that it can retain its lead in parallel processing hardware, the panel recommends that a body of specialists should be brought together to follow the Japanese technical literature to warn of any "technological surprises" coming from the much-vaunted fifth generation computer programme. The British Alvey programme of research and development should similarly be "watched closely". But the panel is ap-

parently unimpressed by the potential of the European Esprit programme.

Other areas where COSEPUP panels urge new federal initiatives include advanced polymer composites, laser-atomic interactions and relativistic plasma waves. Research on polymers is badly organized, according to one panel, and there should be a new emphasis on interdisciplinary training to bring together scientific skills needed for development of advanced organic polymers, possibly concentrated at special centres.

Another panel, on physics, argues the likely future importance of accelerators based on electric fields produced in plasma waves, high resolution spectroscopy using relativistic particle beams, and two-dimensional structures. In plasma physics and astrophysics, there is a need for improved computing facilities.

Australian pests

Legal framework for control

Canberra

THE Australian government's introduction of the Biological Control Bill 1984 is another milestone in the long struggle for supremacy over exotic pests. In the meat export trade, a large part of Australia's reputation as a reliable supplier of high-quality produce is the good health of its farm animals, with no outbreak of foot-and-mouth disease since 1872, rabies since 1867 or rinderpest since 1923.

By contrast, the difficulty in enforcing large-scale biological control of weeds and other crop pests in the Australian federal system is nowhere more clearly displayed than in the drawn-out conflict between beekeeper and landowner over the relative merits of the purple-flowered Mediterranean plant *Echium plantagineum* — known as Salvation Jane or Paterson's Curse, depending on one's point of view, and introduced into Australia more than a hundred years ago.

Since the turn of the century, various grazier groups in the sheep and cattle regions of southern Australia have had the vigorous exotic declared a noxious weed, claiming that it detracts from the fodder value of pasture land and produces pyrolizidine alkaloids poisonous to cattle.

In 1974, all state departments of agriculture responded positively to an inquiry from the Division of Entomology of the Commonwealth Scientific and Industrial Research Organization (CSIRO) to prepare the release of exotic insect agents to control *Echium*. In 1978, supported by graziers from lower rainfall areas, the beekeepers protested that the proposed control programme should not proceed, proclaiming *Echium* the most important honeybee resource growing in south-east

In marked contrast with the recommendations on technology, however, panels reporting on opportunities in basic biology appear content with the existing federal role. An overview of the biology of oncogenes produced by a panel under the chairmanship of Dr Robert Weinberg of Massachusetts Institute of Technology, for example, concludes that further progress towards an understanding of cancer "will depend on continuing support of the basic biologic disciplines whose successes have proved so useful...".

A report on interactions of blood and blood vessels stresses the potential health benefits to be gained and advocates more research but demands no government action. The contrast has not been lost on OSTP. Parasitism, however, is identified as ripe for a major expansion, as molecular biology provides new techniques. Here, as in other areas, OSTP is told of a need to encourage more scientists to study the basic biology of the organisms involved.

Tim Beardsley

Australia and in some years the only spring crop available.

The controversy seesawed for two years until, in 1980, the *Echium* leaf mining moth *Dialectica scallariella* was released at three experimental sites in New South Wales. But later that year, the high court granted a restraining injunction on narrow grounds in favour of beekeeper and low-rainfall-zone grazier plaintiffs. CSIRO was required to halt importation of further European insects and to destroy those already released. Another agreement was reached in June 1982 to set up a tribunal to determine whether control of *Echium* should proceed at all, but this did not materialize.

Throughout the whole litigation period, which ended in June 1983 with CSIRO agreeing to a permanent injunction, the federal government had only a limited capacity to enact far-reaching biological control laws. Except for quarantine, most constitutional power in the area of exotic disease control, prevention and compensation resides with the states.

In August this year, however, the government introduced the Biological Control Bill which has passed in both houses and will soon become law, but only in the Australian Capital Territory. It provides for widespread advertising of proposed control programmes and a commission of inquiry and powers of review for individual cases. The bill in effect negates liability for the approved release of control organisms and is meant as a model for complementary legislation in the states. But because of the inability of either pest or bug to discern state boundaries, opponents of the bill will have to lobby successfully in only one state to disable the entire scheme.

Jeffrey Sellar

European computing

French bid to build centre

TOULOUSE in the south of France — a French centre for the aerospace industries, biotechnology and astrophysics — is making a serious play to become a major European focus for research and training in advanced scientific computation. The aim is to bring together computer scientists interested in new architectures (array or parallel processing, multi-instruction-stream computers and beyond), users and numerical analysts in a melting-pot that would be unique at least in Europe.

The Toulouse centre as at present planned would at least match a combination of the US National Aeronautics and Space Administration (NASA)'s Langley Institute of Computer Applications in Science and Engineering (ICASE) at Hampton, Virginia and the new NASA Research Institute for Advanced Computer Science (RIACS), which started just a few months ago at Ames, California.

This is not, however, a European "fifth generation" artificial intelligence (AI) experiment, although the Toulouse association for the promotion of the Centre Européen de Recherche et de Formation Avancé en Calcul Scientifique (CERFACS) wants to have available the most advanced computer possible, and is aiming to have its real impact in the 1990s.

The CERFACS objective is less ambitious than AI, and possibly more realistic. It stems from a growing feeling among "number crunchers" of all kinds that the new computer architectures and components require radically new techniques, and that interaction between users and computer designers is becoming essential. For example, says Professor Roger Hackney of the Department of Computer Science at the University of Reading, architectures designed to match particular problems can now be "a very cost-effective way of getting large computer power to the problem".

The best use of future and existing machines can be best learned and developed, it is felt, in a centre combining the disciplines of computer science (architectures), numerical analysis (algorithms) and applications specialisms. CERFACS would be such a centre. Nothing like it exists in Europe, Hackney claims. Moreover, NASA's ICASE and RIACS, while addressing all the relevant issues, concentrate architecture on the west coast and applications on the east.

Training and networking would also be "very important" at CERFACS. Visitors would come for several weeks or months, but then return to their home institutions, from which they would wish to keep in touch with CERFACS and use its computing power. So a satellite-based computer communications network is also proposed.

But will CERFACS get off the ground? The idea was the brainchild of Jean-Paul

Zahn, director of the Pic du Midi Observatory near Toulouse, but it has now been taken up with enormous enthusiasm by local industry and politicians such as the local mayor and European parliamentarian Henri Sabi.

However, the news from Brussels is not bright. Although the European Commission has set up a "study" of CERFACS, "we learned last week there's little prospect of Brussels cash in the next five years", said a CERFACS spokeswoman in Toulouse last week. "It just takes that long for the Commission to set up a programme."

So the hunt is on for support from the French government (said to be "very positive" to the idea), and from elsewhere. A European support committee has been established with members from the United

Kingdom, West Germany, the Netherlands, Spain, Greece, Denmark and Switzerland, which will report back early next year with the results of its approaches to government and other agencies.

CERFACS does have one guarantee: a purpose-built building. The Midi-Pyrénées regional government will pay for that. Internationally, Greece has already shown great interest, and smaller European nations may well prove more favourable than, say, Britain and West Germany, which already have large machines — even if they do not use them in the CERFACS manner.

France, however, is well behind Britain and West Germany in scientific computing power, so a conceivable outcome is that France will set up a joint arrangement for CERFACS between itself and the smaller nations of Europe. Whatever emerges from present soundings, a formal proposal to the French government will follow, probably in February. **Robert Walgate**

NIH

Academy recommends guile

Washington

A NATIONAL Academy of Sciences panel, after a year-long study, has suggested that the National Institutes of Health (NIH) might best defend themselves against the pressures of the "disease of the month club" by the time-honoured government strategy of injecting a new layer of bureaucracy into the organization. In recent years, a growing number of disease lobbies, acting through Congress, have pressed for the creation of new NIH institutes bearing the name of "their" disease. NIH has resisted these forces with varying success, but recently scored a victory when President Reagan vetoed a bill that would have established new National Institutes of Arthritis and of Nursing.

The academy recommendation is that there should be a new advisory body, to be called the Health Science Board, whose major purpose, it appears, would be to rechannel the public pressures that now find expression through Congress. The board would "communicate public perceptions of health research needs" to NIH and other Public Health Service agencies and would "assure the public that these needs are being . . . addressed".

Proposals for new NIH institutes would be subject to a formal review by the board, which would apply five criteria that would specifically exclude institutes whose major function is regulation, health services or other non-research activities. By these criteria, the National Institute for Occupational Safety and Health and the National Centers for Health Services Research and for Health Statistics — all of which, at some time, have been considered for incorporation into NIH — would be excluded.

The academy panel would not comment on specific proposals, such as that for an

arthritis institute, but advised that with NIH now having grown to 11 institutes, there should be a "presumption" against new additions. The panel's survey of NIH's organizational and budget history failed to produce evidence to substantiate the most often-heard justification for new institutes — that they boost support for the particular area of research and thus for NIH as a whole. (A spokesman for the Arthritis Foundation, for example, told the panel at a public hearing last year that by so visibly responding to such politically-popular health issues as arthritis, NIH could only gain. "We must see how we can take advantage of the system.")

By contrast, the panel concluded that new institutes add overhead costs, add to the administrative burden of an already overburdened NIH director (25 officials already report directly to him) and hinder scientific communication within NIH.

If the assurances of the Health Science Board are not enough to convince the public that their perceived health research needs are being met, NIH could avail itself of a "continuum" of options short of creating new institutes, the academy panel said. These options include "publicizing what scientific research has accomplished", holding conferences, naming special panels and creating new divisions or programmes within existing institutes.

A somewhat more tangible recommendation of the panel in helping NIH to respond to changing health research needs is for the creation of a contingency fund of 1 per cent of NIH's budget that the director could apply to health emergencies. At current spending, that one per cent equals about \$50 million, which may be more of a slush fund than Congress would be willing to create. **Stephen Budiansky**

Whaling

US and Japan agree to disagree

Tokyo

ANGRY criticism from press and politicians has greeted the Japanese government's decision to bow to US pressure and stop all sperm whaling by 1988. But there is some confusion over exactly what Japan has agreed to, with conflicting announcements in Washington and Tokyo, and protests from Greenpeace (the environmental pressure group) that Japan has got off far too lightly.

According to US negotiators, Japan has agreed to end by 1988 not just sperm whaling but all whaling. With this agreement, the United States hopes to avoid the bad feelings that would arise if, instead, it tried to use cuts in fishing quotas to force Japan



to abide by the rulings of the International Whaling Commission (IWC) that sperm whaling should end this year, and all commercial whaling next year. But Japanese officials claim they have agreed to end only sperm whaling by 1988 and that a ban on whaling altogether is only a US offer. And an angry statement from Hiroya Sanyo, the Fisheries Agency director general who headed the Japanese delegation, adds that the talks had been "unfair" and "outside the bounds of normal diplomacy".

Anti-whaling groups in Western nations are equally angry that Japan is now to be given until 1988 to end sperm whaling. An injunction is to be sought to force the US government to obey the Packwood-Magnusson Amendment to the 1979 Fishery Conservation and Management Act and impose cuts in Japan's fishing quotas the moment Japanese boats catch a sperm whale — and there are unconfirmed reports that two whales have already been caught.

The Packwood-Magnusson Amendment provides that when a nation "diminishes the effectiveness" of an international whaling agreement, it must be "certified" by the US government, which means that that nation's quota in the US 200-mile zone must be cut by at least half in the first year of certification and by 100 per cent in the second year. The outcome of court action is now likely to hinge on the meaning of the words "diminish the effectiveness", with the US government arguing that with the new agreement the effectiveness of the

IWC ruling is not diminished.

Japan is, in any case, far more concerned about minke whales than sperm whales. At the same IWC 1984 meeting where the sperm whale quota was cut to zero, the Antarctic minke whale quota — upon which Japan's offshore whaling industry depends — was cut to 4,224 to be shared between Japan, the Soviet Union and Brazil. Both the Soviet Union and Brazil have formally objected to the quota, thus freeing themselves from obligation to abide by it. Japan has not so far objected because of its fear that US fishery cuts would follow. But immediately after the new agreement with the United States was announced, the agriculture, forestry and fisheries minister Moriyoshi Sato stated

that Japan will file an objection to the minke whale quota before the deadline of 6 January. The following year, the minke whale quota would automatically fall to zero as the total ban on whaling agreed by IWC in 1982 comes into effect.

It seems clear that the minke whale quota has been the subject of a US offer that goes over IWC's head: a blind eye would be turned to further whaling in the next three years and no fishing quota cuts would be imposed, provided Japan will compromise and agree to stop all whaling by 1988. Washington has apparently gone so far as to claim that Japan has already agreed to accept this offer. The subsequent announcement from Mr Sato suggests, however, that Japan will still fight on and try to win some further concessions from the United States — if legal action does not in the meantime undo what has already been achieved.

Alun Anderson

Around the world in four years

OPERATION Raleigh, a four-year circumnavigation of the world that plans to mount over 150 scientific projects in more than 30 countries, was launched last week by its British patron, the Prince of Wales. The flagship of the expedition, named for Sir Walter Raleigh who sponsored the first English-speaking colony in North America 400 years ago, sailed from Kingston upon Hull last week for the Bahamas, on the first of 16 three-month legs.

The expedition is modelled on Operation Drake (1978–80) and is organized by the Scientific Exploration Society Ltd, a registered charity. Operation Raleigh will involve over 300 scientists, assisted by 4,000 selected young people, 75 per cent of whom will come from the United States and Britain. The ship is equipped with three scientific laboratories largely donated by Gallenkamp, diving facilities including a decompression chamber, a Marisat satellite communications system and a computer centre donated by Acorn Computers Ltd and Centronics, and will support both marine and land-based projects.

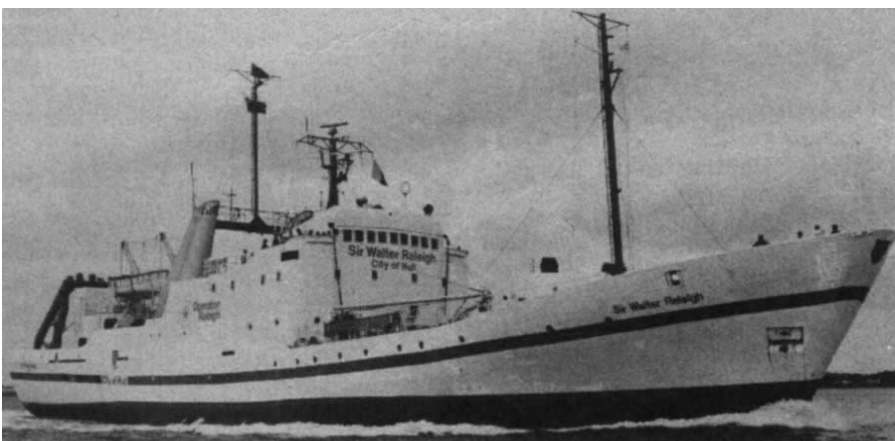
The first leg of the expedition to the Bahamas and Turks and Caicos Islands in the West Indies will include a biological

and geological investigation of the "blue holes" system of the Bahamian islands — underwater caves and tunnels carved out by the dissolution of limestone rocks by fresh water at a time (about 18,000 years ago) when the sea level was about 100 metres lower than it is now.

The Bahamas will also be the first site in a "reefwatch" programme to monitor the effects of sediment pollution on coral reefs. Coastal accumulation of sediment resulting from forest clearance reduces the amount of light penetrating the coral, which has an adverse effect on the photosynthesis of the symbiotic algae on which reef-building corals depend for their survival.

Operation Raleigh, operating as it does mainly from tropical latitudes, might be considered a "jamboree" for the scientists involved, but the main criterion used by the Scientific Research Committee in selecting projects — which must be more or less self-financing — was that the work should produce results suitable for publication in research journals. The young people who will provide semi-skilled assistance on the expedition were also subject to rigorous selection procedures.

Peta Pickering



Vaccinia virus

WHO to back vector vaccines

Washington

A SPECIAL advisory group to the World Health Organization (WHO) has approved an accelerated programme of research on recombinant vaccines using viral vectors. Such vaccines, now under development, might be used in mass immunization programmes where alternatives do not exist or are prohibitively expensive. Though yet to be approved by WHO's regular scientific advisory group, the decision puts to rest speculation that the reintroduction of vaccinia immunization programmes might be thought unnecessarily dangerous.

Vaccinia (cowpox virus) has long been a favourite organism for virologists, being easy to culture and safe to handle. With the elimination of smallpox, its use in vaccination was discontinued, because of severe reactions, sometimes fatal, that occur at the rate of 1 or 2 per million vaccinations. Furthermore, vaccinia can be transmitted by skin contact between people and from animals, albeit rarely.

Despite some reservations expressed at a workshop held last week jointly by WHO, the US Public Health Service and the British National Institute of Biological Standards and Control, the advisory group decided that the potential benefits of research into recombinant vaccines outweigh

the possible risk. Protection against a number of diseases has already been demonstrated in animals using vaccinia as a vector for antigens, including hepatitis B, herpes simplex virus type 1 and rabies (see for example *Nature* 311, 67; 1984). Although the safety of these vaccines has not been systematically studied, it appears that deletion of the thymidine kinase gene of vaccinia — which is replaced by the foreign gene — may make the virus even less easily transmitted than the normal type.

In its advice to WHO, the advisory group stresses the importance of developing alternative vectors likely to be of use against diseases of the respiratory tract and enteric diseases, including bacterial, protozoal and helminthic diseases. Because much early experience will be with animal vaccines, WHO is urged to establish links with the United Nations Food and Agriculture Organization to monitor veterinary results. And to avoid countries using different strains of vaccines haphazardly, WHO is invited to set up a collaborative centre for dissemination of information and storage of reference samples, probably to be located at the National Institute for Allergy and Infectious diseases at Bethesda, Maryland. **Tim Beardsley**

Israeli education

Palestinians push ahead

Rehovot

THE latest of Israel's institutions of higher learning, Ben-Gurion University of the Negev, has finally begun its academic year, two weeks behind schedule. It remains to be seen, however, how long Ben-Gurion and the others — all plagued by severe financial problems — will remain open.

To be sure, Ben-Gurion University, with an accumulated debt of US \$15 million, is worse off than the older, better established institutions. And this presumably explains both the delayed opening and the resignation (later withdrawn) of university president Shlomo Gazit. But with government support for higher education having dropped by one-third in recent years, the other schools are not in much better shape. The president of the Haifa Technion, Professor Josef Singer, has stated that "the Technion is still \$10 million short of even its reduced emergency budget and the conditions for an uninterrupted school year have not yet been guaranteed".

There is greater sympathy for the plight of the universities among members of the present government than existed in the last one, and education minister Yitzhak Navon, for one, has described their situation as "intolerable". But with all public spending being drastically reduced, including substantial cutbacks in even

defence and social welfare, it is unlikely that the government will be more generous than it has been in the past: in all probability, support will decline still further. This has forced the universities not only to undertake large-scale dismissals and reduce their purchases of vital scientific equipment, but also to shut down whole departments.

This development is taking place against the background of a dramatic increase in the educational level of Palestinians, both in Jordan and in the West Bank. During the 1983–84 academic year, 25,000 Jordanians and 13,000 West Bank Arabs passed the Jordanian matriculation examinations, about twice the number of matriculants as among the Jews of Israel, from roughly the same population of some four million people.

Educational advancement has been particularly spectacular in the territories under Israeli control since 1967. From 1967 to 1982, the secondary school population in the West Bank and the Gaza Strip increased by 106 per cent, while the general population went up only 23 per cent; and while in 1967 there were just a few hundred students in post-secondary programmes in those two areas, this year there will be nearly 20,000, a very high proportion.

Nechemia Meyers

Gromyko's prize

THIS year's Soviet State Prizes for science, announced as usual to coincide with the anniversary of the 1917 October Revolution — in November — went almost entirely to "cycles of works" carried out by teams of up to 12 persons over a long period of time. Only four prizes of the fourteen went to monographs by a single author: Ilya N. Vekua's "Some general methods of constructing different variants of the theory of shells" (1982), Dr Andrei D. Ado's "General allergology" (1978), Dr Valentin V. Seldov's "The East Slavs in the sixth to thirteenth centuries" (1982) and foreign minister Andrei A. Gromyko's "External expansion of capital: past and present" (1982).

Several of the "cycles" honoured cover two decades or more including B.B. Bokut' *et al.*'s "Highly effective nonlinear transformations of frequency in crystals and the creation of returnable sources of coherent optical radiation" (1963–82), I.N. Matrosov *et al.*'s "Elaboration of the method of Lyapunov vector functions for the analysis of stability and other dynamic properties of nonlinear systems" (1962–81) and K. P. Belov *et al.* "Magnetism and the electronic structure of rare-earth and uranium compounds" (1959–82). One citation seems to have been overdue — that of N. N. Bogolyubov, D. V. Shirkov and A. A. Logunov for their contribution "Method of renormalized groups in field theory", which carries the date 1955–56.

The citations do not make it clear just how far a cycle of work consists of genuine collaborative undertakings and how far it relates to parallel but separate work on a common theme. Co-winners are frequently scattered throughout several institutes, and even republics. Thus Bokut' *et al.* come from Gomel' University (Byelorussia), the Vavilov State Optical Institute, Moscow State University, the Institute of General Physics of the Soviet Academy of Sciences, Vilnius State University (Lithuania), the Institute of Electronics of the Uzbek Academy of Sciences and the Institute of Applied Physics of the Soviet Academy of Sciences. On the other hand, the team of geologists honoured for their complex study of the Baltic regions of the Soviet Union were all, to judge from their names, native to the Baltic republics (Latvia, Lithuania and Estonia) and for the most part are employed in research institutions of those republics, although one holds an all-Union post, as deputy director of the Marine Scientific Research Association of Engineering Geology. Likewise the team responsible for "Development of the theoretical principles of Soviet criminology" (1961–82) — a unusual subject for a state prize — all seem to be Moscow-based, although affiliated to such diverse institutions as the Academy's Institute of State and Law and the Moscow Police College. **Vera Rich**

United States in UNESCO

Withdrawal sanctioned by default

Washington

AN independent panel appointed to monitor the activities of the United Nations Educational, Social and Cultural Organization (UNESCO) has failed to find persuasive reasons why the United States should reverse its decision to withdraw from the organization at the end of this year. A final decision on US participation in 1985 must be made within the next two weeks, and barring an unexpected change of heart by the State Department, the withdrawal will go ahead as announced at the end of 1983.

The US decision was made because of management inefficiency at UNESCO and "politicization" of its programmes. The monitoring panel, chaired by Dr James Holderman, chairman of the US National Commission on UNESCO and president of the University of South Carolina, makes no specific recommendation on the question of withdrawal. However, its conclusions are thought not to be so compelling as to demand the positive action required to abort withdrawal at this late stage.

UNESCO has made a number of changes over the past year that go some way to answering the US criticisms. The most important, from the US point of view, was the recent agreement to freeze the 1986-87 budget at the 1984-85 level, a move long demanded by the Reagan administration. Furthermore, some of the more ideologically inspired language used to describe UNESCO ambitions on disarmament and human rights has been toned down, and new efforts have been made to evaluate programmes, both actions that indicate some sensitivity to US concern. But recently agreed management reforms are not thought by the United States to go far enough and there is considerable scepticism about how rapidly and effectively the agreed changes will be implemented.

Despite the partial successes, however, the United States has failed totally to bring about changes in the voting structure at UNESCO, and was deserted by its Western allies on this issue. As in other United Nations organizations, there is one vote per country, even though financial contributions are in proportion to countries' gross national products. The US contribution, at \$47 million, amounts to one quarter of UNESCO's total budget. Although UNESCO officials express confidence that programmes will continue even if the withdrawal does go ahead, the impact will be serious — especially if other countries follow the US lead. Britain and West Germany are wavering.

A separate assessment of how UNESCO has accommodated US demands is being undertaken by two subcommittees of the House of Representatives' Foreign Affairs

Committee, under Representatives Gus Yatron and Dan Mica. This assessment will not however be complete before February, and so will not bear on the withdrawal decision. In any event, the executive decision to withdraw would not easily be swayed by a congressional report. At one stage, the possibility was considered that the United States might delay its withdrawal by one year, which would bring it neatly to the end of a UNESCO biennial programme cycle, but that option has now been definitely discounted by Gregory Newell, Assistant Secretary for international organization affairs at the State Department.

The US decision had dismayed many who were working on scientific projects administered by UNESCO, which were thought to be largely innocent of the faults levelled at other programmes. The National Academy of Sciences earlier this year



produced a favourable assessment of the scientific benefits obtained through the agency. But the State Department maintains that in most instances US scientists will be able to continue their involvement with UNESCO programmes even after withdrawal, despite the lack of US representation at council level: domestic expenditure on the Man and the Biosphere programme, for example, is scheduled to be doubled or tripled over the next few years.

The State Department says it is also looking for ways to encourage technological development in developing countries that do not rely on UNESCO. How far these aims prove possible will depend on the seriousness they are accorded by the Office of Management and Budget, and on the willingness of UNESCO to cooperate after being jilted by its biggest paymaster.

Tim Beardsley

West German forests

Damage survey looks bad

Munich

THIS year's inventory of damage to West German forests, the first results of which were published earlier this month, shows further deterioration since last year. A half of all forested areas are now believed to be affected to some degree, compared with 34 per cent in 1983. But as yet there are no reports of forest areas in which the damage is complete.

The Forest Damage Inventory now completed was initiated by the federal government in Bonn, and is the third of a series begun in 1982. For the first time, the survey was based on a nationwide system of 4 km × 14 km plots. This year, however, the classification was refined so as to include five categories of damage (rather than three), numbered 0 (undamaged) to 3 (severely damaged) and 4 (dead).

The main criterion for assessing damage is the extent of loss of needles or leaves, with the discoloration of foliage a secondary consideration. The survey earlier this year was carried out by specially trained forestry personnel carrying colour photographs with which to check their observations.

Between 1983 and 1984, the most obvious change has been the increase of the proportion of West German forests in which the damage is classified as slight or moderately severe. Between the two most recent surveys there has also been a dramatic increase in the extent of damage in some degree to broad-leaved trees, with the percentage of affected beech trees increasing from 26 to 50 per cent between 1983 and 1984, and that of oak trees from 15 to 43 per cent.

In the latest survey, damage has been reported from throughout West Germany, but is most serious in the southern *Länder* of Bavaria and Baden-Wurtemberg. While the earlier surveys showed damage predominantly at high altitudes, the most recent survey shows damage everywhere.

As in previous years, the results of the forest inventory will not go undisputed. Although the methods used in 1984 have been more systematically applied throughout West Germany, the validity of needle loss as the main criterion for assessing damage is likely to remain controversial. But whatever is happening, there can be no dispute that the new type forest decline is increasing both in geographical extent and in severity.

Although the cause of the damage remains open to debate, in the past few years increasing attention has been paid to the importance of ozone in conjunction with other stressing factors as the chief agent of damage. So far, there have been no reports of forest tracts in which all trees have been killed.

L.W.Blank

Conditions for disarmament

SIR — J.N. Barkenbus and A.M. Weinberg (*Nature* 310, 95; 1984) were right to object that Stephen Salter's proposal for a superpowers disarmament negotiations algorithm (308, 490; 1984) assumes an unrealistic degree of collaboration between the protagonists. But it may slightly misleading to term the Salter scheme "original". Undoubtedly it is Mr Salter's own autonomous creation, and in its particular form it may also be entirely novel. It is certainly both elegant and economically structured. However, the "I cut — you choose" algorithm was suggested for this purpose in an article in *Journal of Conflict Resolution* 1963 by E. Singer.

The many differences between the Singer and Salter versions of what the former terms the "divide and choose" approach are probably of little interest, except to those who set more store by the contemplation of abstract negotiation models than I do. Suffice it to notice here that Singer's approach was to place all weapon retentions and weapon discards by both sides into a single population to be divided, and then to consider solutions to the difficulties this could pose. He also throws out the interesting suggestion that on-site inspection rights could be bargained for in the same or another similar structure, after they had first been portioned out in territorial units.

Unlike Salter, Singer did not so much see his scheme as replacing the antagonistic conduct of international relations, but rather as having to take its chances within the sordid arena of diplomacy. Thus he concludes: "Eventually, the realities of international negotiations must be confronted, and the diplomatic foxes must be set free to see if they can, without too much difficulty, play havoc with a bargaining model".

It is a pity that Barkenbus and Weinberg were unable to espouse the realism whose absence they regret in Salter. For it is certainly very naive to suppose, as their model does, that the arms race is about preserving a stable parity between the deliverable portions of populations of militarily undifferentiated "warheads". Their second and probably related assumption, also hard to share, is that we are dealing always and only with retaliatory forces. For once the idea of an attempted disarming first strike is admitted into the strategic calculations (as it always has been on both sides), it is no longer possible to claim that erection of a defensive anti-missile system can be directly offset by dismantling a proportion of one's own offensive missiles, to match the proportion of an enemy missile strike one expects to be able to block. For suppose Side A, with its new defensive system and, say, a matching 10 per cent reduction of offensive forces, were now to initiate a strategic first strike. The 90 per cent offensive force would be virtually as effective as before the reduction, and the defensive force,

facing only the decimated and disrupted remains of B's strategic offensive capability, would be very much more than 10 per cent effective because it would be handling a much smaller salvo of incoming warheads.

Since this is the worst-case scenario which would govern B's perceptions of A's new defensive deployment, it could be argued that the latter should be offset by a reduction of A's offensive arsenal proportional to the hypothetical "thinning" by which B's offensive forces would receive by the two processes of, first, an attempted disarming first strike by A, and second, the anti-missile effects of A's defences on B's subsequent retaliation. Thus A would have to pay for its 10 per cent effective defences with, say, an 80 per cent reduction of its offensive forces.

But of course A could not accept this, because it has to see the world as one in which B could be making the first strike, which demonstrates that there is no way to provide a "fair" mix of offensive and defensive deployments, except by making artificial and unrealistic assumptions to the effect that all nuclear strikes will always be retaliations, which is of course a useless option, despite Barkenbus and Weinberg's so boldly having taken it. R.I.P. BULKELEY
*Department of War Studies,
Kings College London,
Strand, London WC2R 2LS, UK*

SIR — The leading article "Targets for future arms control" (18 October, p.591) was both pessimistic and realistic. In a climate of mutual and reciprocal suspicion between the nuclear armed states, perhaps we should look to small, but achievable, measures that both build confidence and reduce further escalation of the nuclear arms race. I propose a measure that might achieve this and at the same time bolster the Non-Proliferation Treaty (NPT), which, you rightly comment, mercifully still holds, but only tenuously.

With 50,000 or so nuclear warheads worldwide, there is surely no need to add to this total. If so, then an effective "freeze" on fissile material production should be agreed to and implemented by each of the nuclear weapon states. A partial freeze of this sort was agreed by the US, UK and Soviet governments in April 1964.

In order to implement this, the nuclear weapon states would have to open up their reprocessing plants and plutonium/enriched uranium stockpiles to full inspection by a supra-national agency.

In his book *Face the Future* (1981, pp. 488-89), former British Foreign Secretary Dr David Owen has already argued this, basing his conclusions on his experience in office. With the Soviet Union recently accepting International Atomic Energy Agency (IAEA) safeguarding on some of its civil nuclear facilities and with the People's Republic of China joining IAEA,

the foundations of progress exist.

The French and British dual-purpose nuclear facilities, such as Marcoule and Phénix, and Sellafield remain a problem. Indeed Dr D. G. Avery, appearing on behalf of British Nuclear Fuels PLC, told the Sizewell B Inquiry (16 October) that the British Government and Euratom have been at loggerheads since 1973 over the non-accessibility of the Sellafield reprocessing line to physical inspection, despite Britain's accession to the Euratom treaty as a full partner.

Nonetheless, this problem certainly seems surmountable. This being so, it would surely strengthen the NPT, whose review conference in September 1985 is surely going to demand more equitable behaviour by the nuclear weapon states if others are not to withdraw, with the disastrous consequence of a nuclear free-for-all.

DAVID LOWRY

*Energy Research Group,
Open University, Walton Hall,
Milton Keynes MK7 6AA, UK*

Family planning

SIR — The reduction of the birth rate in China by one-half during the 1970s, reported in *Nature* (310, 612; 1984), is indeed welcome news. In view of high levels of fertility persisting in many areas of the world¹, what are the prospects for this feat to be duplicated elsewhere?

The primary requirement for a successful population programme must surely be the participation of most of the population in the programme; yet even in campaigns as vigorous as that in India, fewer than 1 in 4 eligible couples are protected². So the choice of method may be of secondary importance — pills and rhythm, for example, seem to have a similar overall impact on the fertility of acceptors³.

How can the scope of population programmes be broadened in societies which are less highly centralized than China's? A mass-marketing approach may be effective, as demonstrated by the unfortunate popularity of artificial infant feeding formulas in even the poorest areas. A more beneficial twist of this technique might involve the use of some other highly marketable health product in conjunction with a family planning campaign. Due to the threat of anaemia (see for example *Nature* 310, 615; 1984), iron supplements would be beneficial and would have a high level of appeal to women. By distributing supplements in packets of 28 and colour-coding them for the safe and unsafe days of the menstrual cycle, for example, many more women might be introduced to the practice of family planning than have been reached by conventional programmes.

STEVEN L. MARCUS

*Department of Atmospheric Sciences,
University of California, Los Angeles,
Los Angeles, California 90024, USA*

1. McNamara, R.S. *Foreign Affairs* 62, 1107-1131 (1984).
2. Basu, A. M. *Stud. Fam. Plann.* 15, 136-142 (1984).
3. Laing, J. E. *Stud. Fam. Plann.* 15, 49-61 (1984).

Moral responsibility

SIR — In the leading article (*Nature* 25 October, p.692) on the expulsion of the ex-Nazi scientist, Rudolph, from the United States, the statement is made that "scientific research is a neutral, amoral enterprise *only* in a free society, one with a democracy capable of exercising social judgement over the propriety of end uses". I would contend that you let us scientists off too easy, that we all have a moral responsibility, maybe even a duty, not to work on obviously immoral projects, such as, for example, those involving the murder of innocent people by chemical, biological, atomic-nuclear and even "conventional" weapons. Indeed, in the present international situation where warfare is not likely to be between individual soldiers but between peoples of contending states, it is the moral responsibility of scientists everywhere not to work on any project connected with waging war. I know the answer to this: that we in the democratic societies leave ourselves open to "nuclear blackmail". But are you so sure of that, as much as you seem to be unsure of the great possibility that the ludicrous, to say the least, build-up of weaponry will not lead to a nuclear-driven mass destruction?

PHILIP SIEKEVITZ

Rockefeller University,
1230 York Avenue,
New York, New York 10021-6399, USA

Red Sirius

SIR — In the light of the discussions in *Nature* of Homer's wine-dark sea during the past four months, it was almost inevitable that the old problem of the red colour of Sirius would arise sooner or later (*Nature* 6 September, p.8).

However, the problem is soluble, from the testimony of several independent observers.

Arguing for a red colour of Sirius, only two referees may be regarded as reliable: the Roman philosopher Seneca (~70BC)¹ and the Alexandrine astronomer/astrologer Ptolemy (~AD150)². Neither texts, however, gives particulars of how these observations were made. All other Greek and Roman texts often presented as additional evidence³, speak only of the "fires" and "flames" of Sirius as it scorches the Earth during the "Dog-days".

An Assyrian text (BM 118898)⁴, dating from the reign of Ashur-bel-kala (=1070BC), mentions a colour "red as molten copper", but here Sirius is clearly stated to be rising and was therefore observed close to the horizon. The same applies to the observations described by the Arab poet Ibn Alrakka (as preserved by al-Biruni (1000)⁵).

Refuting such evidence are a number of contemporary texts citing a white or blue colour for Sirius. Astrological texts are very useful as stellar and planetary colours were closely observed for inferring future

events. The Greek astrologer Hephaestion Thebanus (~AD400)⁶, citing older Egyptian sources (2BC), discusses the various prognostications derived from the observed colours (golden, red, green and white) of Sirius at its heliacal rising. These colours clearly result from the effects of atmospheric extinction and scintillation on a white star near the horizon. Also, the Roman astrologer Manilius (~AD10)⁷ refers to the cold sea-blue quality of the rays of Sirius.

Most explicit are the observations found in the astronomical chapters of the ancient Chinese dynastical histories. Sima Qian (~90BC)⁸, Ma Xu (~AD100)⁹ and Li Shunfeng (~AD635)¹⁰, discussing the colours of the planets and selected standard stars, all assign to Sirius a white colour.

This evidence clearly shows that Sirius has never been intrinsically red during the past few millennia and is in full accord with the present evolutionary phase of the Sirius-system.

R.H. VAN GENT

Sonnenborgh Observatory,
Zonnenburg 2,
3512 NL Utrecht,
The Netherlands

1. Seneca, Luc. Ann. *Naturales Quaestiones* I, 1.7 (transl. Corcoran, T.H.) *Seneca* Vol. VII (Loeb Classical Library, 1971).
2. Ptolemy, C. *Syntaxis Mathematica* VIII, 1 (transl. Toomer, G.J.) *Ptolemy's Almagest* (Duckworth, London, 1984).
3. See, T.J.J. *Astr. Nachr.* 229, 245 (1927).
4. Grayson, A.K. *Assyrian Royal Inscriptions* Vol. 2, 55 (Harrassowitz, Wiesbaden, 1976).
5. al-Biruni *al-Athar al-Bakiya*, 339 (transl. Sachau, E.C.) *The Chronology of Ancient Nations*, 338 (Allen, London, 1879, reprinted Minerva, Frankfurt, 1969).
6. Hephaestion Thebanus *Apotelesmatica* I, 23 (ed. Pingree, D.) (Teubner, Leipzig, 1973).
7. Manilius, M. *Astronomica* I, 409 (transl. Goold, G.P.) *Manilius* (Loeb Classical Library, 1977).
8. Sima Qian *Shiji* XXVII, 55 (transl. Chavannes, E.) *Les Mémoires Historiques de Se-Ma-Ts'ien* Vol III, 376 (Leroux, Paris, 1898, reprinted Brill, Leiden, 1967).
9. Ma Xu in Ban Gu (ed.) *Hanshu* XXVI, 25ab (ed. Buhzu, Changsha, 1900).
10. Li Shunfeng in Tai Zong (ed.) *Jinshu* XII, 2b (transl. Ho Peng Yoke) *The Astronomical Chapters of the Chin Shu*, 125 (Mouton, Paris, 1966).

Italian slur rebutted

SIR — I was pleased to read in *Nature* (11 October, p.501) a news item wishing well to the new president of the Italian National Research Council (CNR). Indeed, Luigi Rossi Bernardi is a scientist first and a science administrator second, and his previous record is one of achievement through hard work, determination to assess quality and ability to develop rational policies. It is on these objective grounds that his appointment deserves praise. The sensitive reader will perceive this from the report. But it is unfortunate that in his opening paragraph Robert Walgate gives the impression that a major implication of the appointment is a transition from the "underdeveloped south" to the "better organized north". To this the sensitive reader may react emotionally. However, since *Nature* is a

scientific journal, let us look at the issue scientifically.

In volumes 301-306 (1983) of your journal, there are eight papers from Italian institutions. Of these two are from the south, four from central Italy, one from the north. The eighth paper, interestingly, was a collaboration among laboratories in the south, centre and north (Catania, Perugia, Torino). Such a small sample has no claim to statistical significance, but it does not support the notion of a north-south dichotomy with respect to research output, which is what CNR is about. More important, if we look for example to modern biology, it is quite obvious that the main thrust over the past twenty years has come from laboratories in Naples, which, together with other CNR institutes that have developed from them, have today the highest concentration in the country of research and training programmes in molecular genetics and molecular biology.

Nobody will deny the existence of a north-south gradient in economic development in Italy, the historical reasons for which are complex and defy any facile analysis. But it does not become *Nature* to perpetuate ill-founded clichés. The first medical school in the Western world was established in Salerno, near Naples, more than a thousand years ago. Today, among the hardest-working scientists some are Neapolitans working in Naples. One of the challenges facing the new president of CNR is to provide the physical facilities required for a more productive outcome of their efforts and of the talent of their young students. I am convinced that Luigi Rossi Bernardi will be equal to this challenge.

LUCIO LUZZATTO

Department of Haematology,
Royal Postgraduate Medical School,
Ducane Road,
London W12 0HS, UK

SIR — It is unpleasant to read in a scientific journal such as *Nature* an article carrying racist and superficial judgements about Italy and the state of its science ("New Hope for Research Council" *Nature* 11 October, p.501).

Phrases such as "colourful Neapolitan presidency" or "underdeveloped south of Italy" betray profound arrogance and ignorance. You also say that the new CNR president "as a scientist" promises a scientific management of the research council that clashes with his political friendships (a truly scientific asset) and the most incredible statement of all — "... he comes from a medical department, medics being 'the strongest mafia — in the best sense [my italics] in the Italian universities' ". I agree on the strength of that particular mafia. But I am sceptical of any "best sense" underlying that lobby. Mafia is mafia and no benefit can come from it to Italian science.

FRANCESCO IZZO

Casella Postale 87,
I-05018 Orvieto, Italy

Genetic engineering and pharmaceuticals

from John Vane and Pedro Cuatrecasas

Genetic engineering now allows biological synthesis and large-scale production of several proteins with therapeutic potential. The principal challenge in this sphere is to identify new, medically and commercially significant targets — the province of cell biologists, physiologists and biochemists. In the future, genetic engineering will surely provide invaluable tools for the study of the molecular basis of cellular control and pathophysiology, which will permit biochemists and medicinal chemists to design novel medicines.

THE essential ingredients of profitable new ethical medicines are chemicals, which in the past have been extracted from plants (digitalis), animals (insulin, vaccines), living cell cultures (vaccines, interferon) or prepared by the synthetic chemist (most new drugs this century). What lies in the future with the advent of genetic engineering?

Medicines from organisms

The extraction of chemicals from organisms has been a source of important therapies for hundreds of years, giving us some of the most fundamental of today's medicines, including digitalis, curare, ephedrine and atropine. Extraction of the ergot fungus growing on rye led to Henry Dale's discoveries and investigations of acetylcholine, histamine and adrenaline. More recently, the extraction of moulds, initiated by the penicillin saga, has culminated in a whole series of antibiotics, the -mycins and the -sporins, for example. Very recently, avermectins have been discovered in extracts from micro-organisms.

The surface of the Earth continues to be combed for further novel sources of therapeutic agents, although such efforts have become less productive in recent years. In the past 15 years, the search has even turned to the bed of the sea, but the Roche experimental marine station, set up in Australia for that purpose, has now been closed. The US National Cancer Institute's search for natural products led to the discovery of potent plant alkaloids but the programme has been curtailed. Other interesting leads may arise from recent studies of Chinese medicinal preparations. But we may have to accept that the therapeutic natural products have largely been discovered.

How, if that should be the case, shall we maintain the therapies the natural products have provided? Digitalis, the cardiotonic drug, used in medicine in one form or another for at least 200 years and which stimulates the failing heart to beat better, is a good example. The "digoxin" of the 1990s will come from one or more of the following sources:

- The ground-up leaf of the foxglove, as at present.
- The leaves of "super-foxgloves" made to yield more digoxin than before, either by standard plant-breeding techniques or by genetic manipulation.

- A tissue culture technique, based on growing foxglove cells as a slurry or on using immobilized plant cells.
- Chemical synthesis of the complex active chemical in the laboratory and factory, provided that product is cheaper than the three above.
- Replacement of digoxin by a therapeutically superior simpler synthetic chemical that does a similar job, as evidenced by animal and clinical experiments.
- The discovery of an endogenous digoxin-like substance that we all make when healthy, but which is deficient in disease. It is possible that such a substance exists (the natriuretic hormone)¹ and that digoxin stimulates the sites on cells designed to react to this natural substance. The identification of such a substance would provide a different starting point for the synthetic chemist.

Digoxin is not a protein, and the genetic information required for its synthesis in the foxglove is probably so complicated that it is unlikely, in the foreseeable future, that gene transplants into other organisms will lead to cheaper synthesis. Thus, for digoxin, we can safely conclude that extraction of foxgloves (standard, super or slurry) will continue until such time as the medicinal and organic chemists provide better — cheaper, less toxic, more efficacious — alternatives. But that is likely to happen during the next decade.

The same general conclusion can be extended to other natural products, bleomycin or adriamycin, for example, for the skills and power of the medicinal chemists have now outstripped nature. Computer chemistry will speed this process.

Animal sources

Other natural medicines are extracted from animal tissues or animal cell cultures. Two kind of extracts are possible. The first is typified by insulin, where a useful chemical is produced from the glands of animals (cows or pigs) which are bred and killed for other purposes. Thus, insulin is a valuable by-product of a different industry. Fractions of different components are also made from human blood. A more macabre example is the extraction of human growth hormone from the glands of dead people. The second kind of extract is the basis of the vaccine industry, which has tradition-

ally relied on growing bacteria or viruses in enormous quantities, treating them in such a way that their pathological activity is lost but their capacity to immunize (immunogenicity) retained. When the resultant "soup" is injected into animal or man, the immune system reacts defensively to the antigenic content (usually a protein or carbohydrate) in such a way that the whole organism is prepared to defend itself against invasion by the living bacteria or virus. The soup sometimes contains live attenuated viruses which lead to actual viral replication in the host, thus eliciting a productive antibody response (as in mumps, measles, rubella, polio, yellow fever); such vaccines could not easily be supplanted by synthetic approaches, and they exhibit potent adjuvanticity and long-term protection.

Somewhere between the two examples are substances like interferon, interleukins, urokinase and tissue plasminogen activator with actual or potential therapeutic actions. These can be produced in quantity from human cells grown *in vitro*. For example, human α -interferon can be produced in commercial quantities by deep cell culture methods pioneered by The Wellcome Foundation.

Traditionally prepared vaccines would never meet the standards required of synthetic medicines, with the demand that the inactive ingredients should be characterized and that the active principle should be more than 90 per cent (and sometimes more than 99 per cent) pure. The antigen or active ingredient of a vaccine may represent less than 1 per cent of the material in a mixture of hundreds of unknown by-products of the bacterial, viral and animal material used to generate them.

It is in this overall field of protein chemistry that genetic engineering offers to provide directly applicable products of therapeutic value and commercial reward. Large protein molecules are, at present, too complex for the synthetic chemist to tackle, so that the biosynthetic route offered by gene-splicing is enormously promising for the large-scale production of these proteins. Furthermore, proteins of animal origin often differ slightly in chemical structure from those of human origin, so that when given to man for replacement or other therapeutic use, they may be recognized as "foreign" and thus lead to unwanted reactions. For this reason, an additional attraction of the

gene-splicing route is that it can provide human proteins. Many of these, such as insulin, interferons and growth hormone, have already been successfully prepared and are either on the market or in clinical trial. The major challenges lie within a limited number of potential candidates from which only a few may offer significant medical value and commercial potential.

Gene products

We are now seeing the commercialization of insulin and interferons derived by gene-splicing techniques. Human insulin is already on the market. However, it is still not clear whether the investment that has been made to fulfill this specific purpose will ever be justified by sales of human insulin alone. Immunogenic intolerance to bovine or porcine insulin (requiring treatment with human material) is extremely rare, the synthetic insulin appears to produce immunological problems of its own and, in any case, the present pancreas-extracted insulins are cheap. The extraction from the pancreas of animals is not only a side-product of the meat industry, but has also become a multifaceted industry in its own right, yielding enzymes as well as insulin; until the enzymes as well as the insulin are more cheaply available from gene-splicing techniques, the very slight differences in side effects predicted for human insulin may not provide the necessary commercial edge. Moreover, some of the goals of genetically engineered "new" or different proteins can also be achieved by chemical modification (by traditional techniques) of available proteins (such as bovine insulin). Human insulin derived from gene splicing may turn out to be a "loss leader".

Biosynthetic interferons could well be different. There are aspirations that if a gene-spliced interferon can be made pure and commercially viable, it may be used in conditions such as the common cold² and for other viral diseases, including influenza, although present experience suggests that the side effects may not be acceptable for most of the self-limiting and mild viral diseases. Given that more refractory medical conditions (such as hepatitis and solid tumours) respond to interferons, they may be particularly attractive targets for, unlike insulin, (1) they are not available in any single tissue or organ for simple extraction (they are not stored in tissues); (2), they have proved extremely difficult to purify in significant quantities; (3) there are several chemically distinct classes (and probably subclasses) of the hormone and, (4), large amounts are needed for clinical experimentation simply to determine their specific utility.

Other human proteins also in the sights of the genetic engineers include growth hormone (but will it be commercially viable?), serum albumin, urokinase, plasminogen activator, factor VIII and lymphokines (the potential applications of

which are presently unknown). Most such potential medicines will need to be injected because of their chemical nature. Drugs available orally will always have more appeal than drugs given by other routes. Simpler but equally effective small molecules capable of synthesis by chemists will always be more attractive than complex proteins, no matter how synthesized. It is for these (and other related) reasons that the biosynthetic route, including gene-splicing, will always be threatened by advances in biology and biochemistry and by chemical techniques. In fact in most spheres of activity, the genetic engineering approach to the biosynthesis of specific molecules is a prelude to (and will encourage) development of superior and less expensive alternatives by organic chemists and biochemists.

The genetic engineering approach, coupled with Milstein's monoclonal antibodies³ and affinity methods for protein purification⁴, already offers an alternative method for the characterization of proteinaceous hormones, transmitters and neuropeptides of potential therapeutic importance. Many such hormones (enkephalins, endorphins, interleukins, morphine, lymphokines, transforming peptides, macrocortin) have been recognized in the past few years, and there are clearly many more to come. The tools of the genetic engineer will provide reagents that will contribute greatly to this field. Not only will the new techniques provide large quantities of pure peptides for study, but they will also assist in identifying and isolating the targets, or receptors, of their action. New and important therapeutic avenues must arise from this work. Recombinant genetic approaches will be invaluable for the study of cellular control mechanisms. Furthermore, recent advances in understanding the nature, chemical structure and function of numerous receptors provide essential information for the medicinal chemist. The ultimate therapeutic value of highly active molecules such as interferon, insulin, growth factors and lymphokines lies not so much in their end-use as therapeutic agents, but more in what they teach us about the disease process. It is in this context of new knowledge and understanding that truly selective and rationally derived new therapeutic agents will arise. Thus, understanding of the mode of action of peptide transforming growth factors, and the expression and target activity of oncogenes, may ultimately lead to the selective control of the malignant state.

Vaccines

Genetic engineering offers many opportunities for the development of new and improved vaccines. Some, like foot and mouth disease vaccine (FMDV) and hepatitis, are near at hand. Others, like a human malaria vaccine, are within sight⁵. The targets range from an attack on present markets such as FMDV through a range of

achievable new human and animal vaccines against bacterial, viral and protozoal infections, to still hypothetical vaccines that might make sheep the size of elephants, or prevent pregnancy. Those who side with prophylactic rather than therapeutic medicine can only be strengthened by the reality of gene-splicing. The progress that may be expected in the vaccine field is shown in Fig. 1. We are through phases A and B and into phase C. To understand the further progress which is possible needs only an elementary understanding of protein chemistry. The amino acids from which proteins are strung together are each about the same chemical size (molecular weight about 150) as one of Dale's transmitters, such as histamine or adrenaline. While the smallest of the proteins (or peptides) of 1 to 50 amino acids, such as angiotensin, may retain a linear structure, larger proteins (several hundred amino acids) wrap up into a three-dimensional complex, each with its own geographical characteristics determined by the type and number of amino acids.

Although it was originally thought that the ability of foreign proteins to stimulate the body to react defensively by producing antibodies (their immunogenicity) was linked primarily to the shape or tertiary structure of the whole molecule, studies with plant and bacterial viruses have shown that antiviral activity can be elicited by a small synthetic peptide corresponding to a region of the antigen involved in virus neutralization⁶. This work has been extended to animal viruses to induce an immune response to FMDV using a small linear part (14 amino acids) of the protein antigen⁷. This suggests that "a small peptide may be able to adopt a conformation approximating that which it assumes in the virus particle". Similarly, total synthesis of a diphtheria vaccine based on fewer than 20 amino acid residues has been achieved⁸.

Whatever may be the conformation needed to elicit antibody activity with such peptides, these results imply enormous consequences for the production of antigens by gene splicing techniques. For a peptide consisting of a score of amino acids is well within the scope of synthesis by medicinal chemists, and so is potentially a synthetic chemical rather than a biosynthetic chemical or a "biological". Clearly, should this work be predictive for other single protein vaccines, the transition from phase C to phase D (Fig. 1) could well take place in the next few years.

Ten years ago, the idea of a peptide carrying activity into the body via the oral route was totally unrealistic, simply because the gut is designed to break down and provide a protective impermeable barrier for larger molecules (with, say, molecular weight exceeding 1,500). Now, orally active molecules based on a peptide structure are becoming common, through the skilful insertion of "unnatural" substituents. Two examples will illustrate

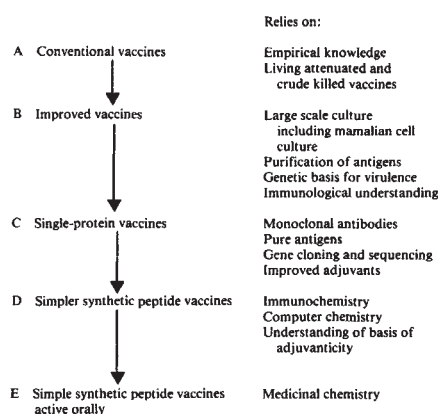


Fig. 1 The pattern of progress in vaccine production.

the point. The first is the development of the drug Captopril, an antihypertensive agent based upon inhibition of angiotensin converting enzyme. The first such inhibitors were peptides extracted from snake venom⁹. Knowledge of their structures led to the synthesis¹⁰ of a hybrid molecule which retained the required activity, but was orally active. The second example is in the field of enkephalins, where replacement of the naturally occurring L-amino acids (such as alanine) in the pentapeptide by the unnatural D-isomer¹¹ conferred a stability against enzyme degradation which allowed oral activity.

Thus, the transition from phase D to phase E (Fig. 1) is already theoretically possible. But a serious limitation to progress (even from B to C) is the need to strengthen the body's immunological recognition of and reaction to antigens by the use of adjuvants. This is easier to do in animals because adjuvants of a more toxic (or inflammatory) nature can be used. In man, anything worse than a sore arm for a few days would be unacceptable. For this reason, it is essential that better adjuvants be found. Perhaps even more important, we need to understand the mechanism of action of adjuvants in order to go to phase E. This has, surprisingly, been a generally refractory field of research.

Chemical synthesis

Virtually all new chemical entities of major therapeutic importance in the immediate past (such as propranolol, allopurinol, acyclovir, cimetidine) have arisen from chemical synthesis, and will do so in the future. There have recently been stupendous advances in biology and medicine, and numerous new molecular targets (enzymes, receptors, transmitters, mediators) are presently being identified as attractive and rational therapeutic points of leverage. At the same time, the ancillary tools are becoming ever more powerful. Furthermore, the automation of chemical reactions is allowing the synthesis of molecules of a type, and on a scale, unheard of even 20 years ago. This is especially true of the peptide field. On the one hand, the synthetic chemist is able to put together longer peptides, on the other,

the peptide needed to convey a specific antigenicity as a vaccine is getting smaller. It seems inevitable, therefore, that the medicinal chemist will eventually dominate prophylaxis as well as therapeutics.

Another enabling method that reinforces this conclusion is the computer chemistry or computer graphic technique¹². By visualizing the actual tertiary structure of receptors (also proteins), and the ways in which drugs interact with their active sites, the rate at which the chemist will be able deliberately to design heterocyclic molecules to fit into or modify such sites will be increased dramatically. These methods should yield standard principles applicable to the synthesis of unique custom-made molecules. For example, sufficient knowledge regarding the manner by which insulin (or interferon) interacts with its receptor at the molecular level should lead chemists to the design and synthesis of simpler, yet specific small molecules, not necessarily peptides, which will do the same (or a better) job as insulin (or interferon). The complex biological molecules will then no longer be needed as medicines.

Drug delivery

A new chemical entity, whether synthetic or biosynthetic, not only has to be active but also has to reach that part of the body where it is needed. Where an illness is associated with a particular organ (such as the liver or lungs), or with a discrete site (such as a solid tumour), it is advantageous to try to concentrate the drug in that tissue needing treatment, so that adverse effects on other tissues are minimized.

Simple ways of localizing or concentrating drugs have been known for many years. The inhalation of anti-asthmatic drugs theoretically takes substances directly to the organ concerned. A more exotic way of delivering drugs to a particular site or organ is to tie them to a messenger addressed to that organ. Antibodies directed at particular tissues have an obvious potential, and many groups are now searching for monoclonal antibodies that will attach specifically to cancer cells. A general antibody that reacts with cancer cells but not with normal cells will have not only diagnostic application but also the potential to carry a toxic message directed only at cancer cells. A further advance (if not a requirement, for selectivity) will be to direct the messengers to molecular entities on the cell surface which are normally destined to be internalized, or directly to the interior of selected cells, where they can do their damage.

This is an area of development in the biotechnology field of substantial importance. It is unlikely that, in the foreseeable future, medicinal chemists will find better messengers than can the immunologists. But the skills of the medicinal chemist will be needed to put the message into the hands of the specific chemical messengers and even to devise the

nature of the message, which may be radiochemical.

Future prospects

Over the next 20 years the various techniques associated with "genetic engineering" will infiltrate all fields of industry and science. Agriculture, energy supply and farming will all become more efficient. In science in general, and in those sciences applicable to the pharmaceutical industry in particular, there will be a substantial acceleration in the acquisition of knowledge of the working of different cells, and of the ways in which they communicate with each other. Monoclonal antibodies and gene cloning will also contribute significantly to the diagnosis and correction of genetic disorders. Importantly, the very knowledge which will be generated by the techniques will enable other disciplines, such as pharmacology, biochemistry and chemistry, to reap the fruits of progress.

At the same time, however, the synthetic chemical routes will eventually replace the biosynthetic routes offered by gene-splicing. For example, breakthroughs with synthetic chemical antiviral agents, such as Wellcome's new antiherpes drug, acyclovir¹³, will be followed by others that could make interferon obsolete. Similarly, insulin will be supplanted when there are discovered small molecules which mimic its action (as opioids, enkephalins and endorphins can resemble each other at receptor sites). Other simpler molecules will surely replace macrocortin, interleukins, endorphins, peptide hormones and so on.

From all this, it should be evident that the thrust of innovation depends on many disciplines working together. Creative biology and medicine must be the enabling force, by providing a body of knowledge which can be exploited by chemistry. In addition, the skills of the genetic engineer and the immunologist are essential not only to provide new (if temporary) prophylaxis and therapy but, much more importantly, to contribute to fundamental knowledge. For this reason, they should be integrated as closely as possible into the existing patterns of discovery rather than be isolated, as means to commercial ends, in their own compartments. □

1. Gruber, K. A., Whitaker, J. M. & Buckalew, V. M. *Nature* **287**, 743-745, (1980).
2. Scott, G. M. et al. *Lancet* **ii**, 186-188, 1982.
3. Milstein, C. *Scient. Am.* **243** (4), 66-74 (1980).
4. Cuatrecasas, P. in *Advances in Enzymology* (ed. Meister, A.) 29-89 (Wiley, New York, 1972).
5. *Bull. Wild Hlth Org.* **61**, 81-92 (1983).
6. Arnon, R. A. *Rev. Microbiol.* **34**, 593-618 (1980).
7. Bittle, J. L. et al. *Nature* **298**, 30 (1982).
8. Audibert, F. et al. *Nature* **289**, 593 (1981).
9. Ferreira, S. H. *Br. J. Pharmacol.* **24**, 163 (1965).
10. Cushman, D. W. et al. in *Enzyme Inhibitors as Drugs* (ed. Sandler, M.) 231-247 (Macmillan, London, 1980).
11. Beddell, C. R. et al. *Proc. R. Soc. B* **198**, 249-265 (1977).
12. Haggin, J. *Chem. Engng News* 9 May, 7-20 (1983).
13. Ellison, G. B. *Am. J. Med.* **73**, 7-13 (1982).

John Vane is at Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, UK and Pedro Cuatrecasas is at Wellcome Research Laboratories, Research Triangle Park, North Carolina 27709, USA.

Who will clone a chromosome?

The cloning of factor VIII, on which most haemophiliacs depend, is a technical triumph without parallel, but has left important questions unanswered. They rest for others.

By sad coincidence, this week there may be special interest in the publication of four long articles (pp.326 *et seq.*) describing how two groups in the United States have been able to isolate the gene responsible for human factor VIII, one of the essential components of the blood-clotting system. The significance of this development is described on p.326 by Professor George Brownlee and Dr Charles Rizza.

They explain that Genentech and the Genetics Institute have been able to clone the gene for factor VIII, to work out its nucleotide sequence, to infer the way in which the corresponding very large protein molecule is put together from amino acids and even to devise a system by means of which factor VIII may eventually be manufactured in large quantities.

This is also the week in which people in Australia have been startled by the news of the death of at least three infants as a consequence of transfusion with blood now suspected of being contaminated with the virus (called HTLVIII) responsible for acquired immune deficiency syndrome (AIDS), and in which a British haemophilia patient died from what appear to have been similar causes, the contamination of supposedly pure factor VIII by the same virus.

Although people in the United States have been familiar with such tragic happenings for at least two years, the spread of this bad news is a chilling reminder both of the dangers to us all embodied by the virus responsible for AIDS and of the urgency of the need for a process for producing factor VIII that may be free from risks of that kind. It is especially cruel that people suffering from haemophilia, already a sufficient hazard to peace of mind, should now be exposed to a further risk over which they have literally no control. Whatever the commercial interests of those responsible for this week's reports, their fleetness of foot will earn full-throated congratulation. Such is the promise of what we call genetic engineering.

Unfortunately, this week's clutch of papers will also help to give that field a less enviable reputation. In the application of molecular biology to practical affairs, it is for the time being clear that bigger battalions move faster. So is there a danger, already explicit in high-energy physics, that the list of the names of the authors of a scientific article will be longer than a reasonably succinct listing of its

conclusions? Readers must judge for themselves, recognizing as they do so the extraordinary character of what has been accomplished.

Hitherto, even the natural site at which factor VIII is synthesized has not been known (but the liver was always a candidate), the protein molecule is the largest whose genetic origin has yet been characterized and the circumstance that the protein (but not, yet, the glycosylation parts that naturally go with it) can actually be manufactured on a laboratory scale would not have been thought possible even at the beginning of this year.

Two sceptics' questions nevertheless arise, the first of which is whether, in the single-minded pursuit of something of undoubted value, those responsible for this staggering exploit may have neglected tantalizing scientific issues that could in the end turn out to be more interesting. And, perhaps, even more important.

Here is an example. To many readers, the most arresting feature of the stack of papers now published will be the information, mentioned almost in passing by both groups, that part of the factor VIII molecule now synthesized resembles in its amino acid sequence part of the molecule of a quite unrelated protein, the plasma protein called ceruloplasmin, whose function is for the time being largely unknown. Ceruloplasmin has two unambiguous properties — it is blue in colour (which is what its name implies) and it binds copper atoms (between four and six per molecule). But properly supplied with copper, ceruloplasmin is an oxidizing enzyme. The common guess is that its function in the human plasma is to convert ferrous (Fe^{2+}) into ferric (Fe^{3+}) ions. But nobody seems to have conclusive proof.

So why, teleologically speaking, should factor VIII resemble such a molecule? There are obviously several candidate answers. Perhaps factor VIII, like ceruloplasmin, also binds copper atoms in ways that are important to its function. (Molecules of factor V, the analogous component of the blood-clotting system, have at least one copper-binding site.) As yet, it seems, nobody has checked.

It is also, however, possible that evolution, always ingenious, may have hit on a way of using a pre-existing molecular structure (ceruloplasmin) as a way of helping blood to clot rather than its normal function, for the sake of argument that defying the second law of thermo-

dynamics, by banishing ferrous ions when they would normally occur. (Ferrous ions are an obvious threat to living things; ferrous chloride is virtually insoluble, they form free radicals easily.) The trouble is that is nobody, even now (and after reading pp.326 *et seq.*) can know whether factor VIII is a copper binding protein. Factor V is, but that could be a different matter.

In normal times, if the field were stagnant, these speculations would have been suitable material for a successful grant application (or several). So too would have been some of the other questions not merely unanswered but unasked by pp.326 *et seq.* Why (or, rather, how) does it come about that such a basic system as that responsible for the clotting of blood should require not one enzyme but several? (Because evolution has found it to be worthwhile, at the margin as the economists would say, to win a little extra efficiency at the cost of a great deal of extra elaboration, perhaps because the molecules required for these marginal purposes were already in essence available.) Why are the components of the blood-clotting system distributed so widely on the human chromosomes? (The same answer applies; they are where they happened to be, so to speak.)

But how does factor VIII function? What is the mechanism by which, in human plasma (and in association with another shadowy plasma protein, von Willebrand's factor) it catalyses the transformation of factor Xa? Nobody has begun to guess what may be at the bottom of this process. Nobody can be blamed, for even a few months ago the question would have seemed insufferably impertinent. Yet it needs to be answered, and quickly.

That raises the second question about the balance of this triumphant effort in the cloning of factor VIII that may yet come round to haunt the cloners — is it feasible to think of manufacturing by the kilo a molecule as large as that of factor VIII without risking that it will emerge corrupted, or as bad, ineffective, from the production line? Obviously, the genetic engineers would prefer to be manufacturing a simpler molecule, one related to the molecule whose structure had now been described, as a few amino acids in an antigenic molecule may represent the biological activity of the entire structure. Will that be feasible? Time, at the worst, will tell. Otherwise, a little research may make the question answerable.

John Maddox

Biotechnology

Clotting factor VIII cloned

from George G. Brownlee and Charles Rizza

ONE of the most exciting and ambitious goals of the biotechnology industry has been achieved. In papers that begin on page 326 of this issue, both Genentech, assisted by Tuddenham and colleagues of the Royal Free Hospital, London, and Genetics Institute, assisted by Fass and Knutson of the Mayo Clinic, Rochester, report the cloning of DNA for human blood clotting factor VIII:C and the production of the factor from cloned DNA expressed in mammalian cell lines.

Clotting factor VIII:C is the name given to the protein that is absent or defective in patients with classical haemophilia (haemophilia A), a sex-linked bleeding disorder of about 20 per 100,000 males. In the blood stream, factor VIII:C is associated with another protein, known as von Willebrand protein or factor VIII-related antigen, but factor VIII:C is solely responsible for the clotting activity of the complex (which is loosely termed factor VIII). Injection of factor VIII obtained from donor blood allows haemophiliacs to be successfully managed. But, unfortunately, the factor VIII preparations in use at the moment carry a high risk of transmitting blood-borne viruses such as non-A, non-B hepatitis virus. The virus that is thought to cause AIDS may also contaminate preparations and, although relatively few haemophiliacs have been affected so far, the fear of the disease has caused great anxiety to patients, their families and their doctors. A genetically-engineered factor VIII:C derived from a source other than blood should be free of such viruses. It should also be cheaper.

Both groups have used similar methods for the DNA cloning and expression of factor VIII:C and report very similar results. Oligonucleotide probes were synthesized on the basis of small sections of the amino acid sequence of purified factor VIII:C (either pig or human in origin), and used to screen an appropriate library of genomic clones to identify part of the factor VIII:C gene. The gene itself turned out to be extraordinarily long (180 kilobases), to have 26 exons and to occupy an estimated 0.1 per cent of the X chromosome (see Gitschier *et al.* on page 326 for further details).

For expression of protein it was necessary to isolate cDNA clones. These were obtained either from a human T-cell hybridoma line (Genentech) or from human liver (Genetics Institute) and sequence analyses allowed the derivation of the entire 2,332 amino acid sequences of the mature protein. Both groups (but see Vehar *et al.* on page 337 for an extended discussion), note that the protein is highly glycosylated and has an obvious domain

structure with predictable homology to another clotting factor, factor V, but also an unpredicted homology to ceruloplasmin — a serum protein believed to be involved in copper ion transport in the blood. The protein sequence also allows the definition of the main sites that are cleaved by thrombin in the process of activation of factor VIII:C.

The expression of factor VIII:C in mammalian cells required the reconstruction of a full length cDNA clone and its attachment to a viral promoter sequence. When this construction was introduced into either a hamster kidney cell line (Genentech) or a monkey kidney cell line (Genetics Institute), a human factor VIII:C-like activity was secreted into the media in which the cells were grown. The concentrations of factor VIII:C-like activity in the media were about one per cent (Genetics Institute) or seven per cent (Genentech) of the normal plasma concentration as assayed by a sensitive biochemical test based on the activation of factor X and hydrolysis of a chromogenic substrate. Furthermore, both groups showed that the secreted activity (further purified by affinity chromatography at Genentech) was able to correct the clotting time of plasma from a haemophiliac. Evidence that these assays are indeed a measure of factor VIII:C and not of some other clotting factor or non-specific activity, was provided by two additional tests carried out by both groups. In the first, the clotting activity was shown to be increased by

thrombin, a normal physiological activator of factor VIII:C. In the second, it was shown to be inhibited by specific factor VIII:C antibodies. The Genentech group also showed that, like authentic factor VIII:C, their substance bound reversibly to the von Willebrand protein immobilized on an affinity column and that its activity was destroyed by protein C — a serine protease of plasma that can inactivate factor VIII:C.

What is the outlook for treatment of haemophiliacs? It is clear from previous experience with genetically-engineered products that we must anticipate some delay, perhaps 3–5 years, before a safe, effective and clinically-tested product is on the market. The expression of factor VIII:C reported by both groups is very low and it is probable that more effective vectors or host-cell systems will have to be devised. Effective large-scale tissue culture and purification methods will also have to be developed. These may differ from existing procedures for purification of factor VIII from blood. We emphasize the need for a high purity product to guard against the possibility of any pathogenic material from the culture being introduced into haemophiliacs. It would be unacceptable to eliminate the risk of AIDS but substitute a new hazard. There remains the possibility that the genetically-engineered factor VIII:C, although of proven activity in a test-tube assay, may not be as effective *in vivo*. This can only be adequately tested by clinical trials of the product, the outcome of which must be eagerly awaited. □

George G. Brownlee is in the Sir William Dunn School of Pathology, South Parks Road, Oxford, OX1 3RE and Charles Rizza is director of the Oxford Haemophilia Centre, Churchill Hospital, Oxford OX3 7JL, UK.

Astrophysics

Evolution of star-planet systems

from Geoffrey Bath

A PROBLEM at the heart of stellar astrophysics is the role of angular momentum in the evolution of stars. This statement may seem controversial, given the profound success of stellar evolution theory in describing quantitatively the nucleosynthesis of the elements and the evolutionary transformation in the structure of stars as they perform their function as cosmic alchemists. The theory of stellar evolution has been largely developed without involving rotation in any major way. But a recent paper by M. Livio and N. Soker (*Mon. Not. R. astr. Soc.* **208**, 763; 1984) raises the dramatic possibility that the evolutionary fate of many single stars accompanied by planets could be both binary star systems and exploding variables.

Certainly, in the case of the Sun, present angular momentum, if rotation is uniform,

is quite unimportant in determining structure or immediate evolutionary future. The Sun cannot, however, be considered as an isolated unit; it is only one part — albeit by far the most massive part — of the Solar System. The angular momentum of the planets is sixty times greater than that of the Sun's own, assuming solid-body rotation in the inner core. Furthermore, the initial solar nebula, with a minimum mass of some 3 per cent of the Sun's mass, must have possessed an even greater fraction of angular momentum.

In the case of binary stars, angular momentum plays a very important part in determining their present binary state and subsequent evolution. In binaries, angular momentum decreases with decreasing separation. Even for W Ursae Majoris

stars, however, in which two relatively low-mass and approximately main-sequence stars are orbiting in physical contact, exchanging energy and surrounded by a common envelope, the angular momentum is hundreds of times greater than that of the Sun. Since more than one-half of all stars in the Galaxy are binaries, then if potential planetary systems around single stars are ignored, the significance of angular momentum is evident.

It was angular momentum considerations of these sorts that led Otto Struve in the 1940s to make the radical hypothesis of an evolutionary sequence that goes from giant binaries to dwarf binaries, and then to single stars with accompanying planets. He envisaged this being a result of the progressive dissipation of energy and loss of angular momentum associated with turbulence and differential rotation that von Weizsäcker was then advocating. As a general evolutionary theory Struve's notion has to be rejected, largely on the grounds that the dissipation time scales are too long in general to compete with nucleosynthesis as the main process of stellar ageing. There is incontrovertible evidence, however, that at least some stages of Struve's sequence have been significant and account for the formation of at least one important class of stars, the cataclysmic variables — though how is still a matter of debate.

Cataclysmic variables are binaries composed of a main-sequence star and white dwarf companion. The stars are so close that material from the surface layers of the distorted, pear-shaped, but otherwise apparently normal main-sequence star is being gravitationally captured by the white dwarf. Cataclysmic eruptions characterize these stars as novae and related exploding stars, generated both by unstable bursts in the mass transfer process and through runaway nuclear fusion reactions in the accreted hydrogen-rich material on the white dwarf surface.

One outstanding problem is the origin of these variables. Why is a white dwarf, which is formed in the deep interior of a red giant, now so close to a main-sequence star? In order to allow its precursor giant to form, the separation there must have been > 200 solar radii ($R_{\odot}$), whereas now it is $\approx 2 R_{\odot}$. Evidently either angular momentum has been lost, and lost relatively rapidly, or an extreme mass ratio originally allowed large separation with the present quota of angular momentum.

Livio and Soker have been exploring in some detail the second hypothesis, as originally advanced by P. Eggleton. They have examined the interaction of a planet immersed within a red giant's extended envelope. As it orbits within the envelope, tides are raised, viscous drag leads to dissipation and loss of orbital energy, and in certain conditions mass may be accreted efficiently. As a result, the planet acts as a catalyst, spiralling in through the envelope and accreting some fraction of the stellar

material as it does so. Livio and Soker find that the final mass of the planet is almost independent of the initial planet's 'seed' mass and is $\approx 0.14 M_{\odot}$, whereas the separation falls from $400 R_{\odot}$ to the range $8.6\text{--}0.5 R_{\odot}$. In this model for the creation of cataclysmic variables, loss of angular momentum is not invoked. Rather it is the change from the extreme mass ratio of a planet-giant system to the less extreme dwarf-white dwarf system that drives the two components together.

The present treatment is limited to somewhat *ad hoc* envelope conditions and assumptions of angular momentum conservation. Nonetheless, the principles by which a planet or intermediate dwarf may be converted into a low-mass main-sequence star, via the stripping of a red giant envelope to expose the degenerate core, are clearly revealed. More recent work by the same group shows that the

same features are present when the giant envelope reacts self-consistently to the spiralling process. If this is proved by future observational tests to be the origin of cataclysmic variables, Struve's ideas will have found some support, but with the important difference that the evolution from a wide planetary system to close interacting binary is in the opposite sense to Struve's early thesis. But then stellar evolution of single stars was long thought, erroneously, to proceed from red giant to main-sequence dwarf. With the idea that the fate of our own Solar System might conceivably lie as a cataclysmic nova, a new possible future for the Sun has to be confronted. □

Geoffrey Bath is at Wolfson College and is in the Department of Astrophysics, University of Oxford, Oxford OX1 3RQ, UK.

Molecular biology

Multiple levels of gene control in eukaryotic cells

from Geoffrey North

How can a gene be regulated by steroid hormones in some cells of a chicken and not in others, in which it is either expressed constitutively or not at all, even though both categories of cells may contain perfectly functional steroid hormone receptors? And how is the level of expression of immunoglobulin heavy-chain genes increased approximately 30-fold when B lymphocytes differentiate into antibody-secreting plasma cells? The possible answers to these puzzles, which emerged at a recent meeting*, illustrate how multiple levels of control combine to ensure that the right genes of eukaryotes are expressed at the right times, in the right amounts and in the appropriate cells.

The pattern of immunoglobulin heavy-chain gene expression in B cells has been presumed to be at least partially accounted for by the 'enhancers' that were discovered last year between the joining (J) and first constant (C_{μ}) exons of both human and mouse genes¹⁻³. These were the first gene-regulatory sequences of eukaryotes to be discovered that manifest cell-type specificity. Typically for enhancers, they can stimulate transcription of a gene whichever orientation they are in and at variable distances upstream or downstream from the gene. However, as W. Schaffner (Institute für Molekularbiologie, Zürich) explained in Heidelberg, this clearly cannot be the whole story.

The original evidence that these enhancers are cell-type specific came from experiments in which they were linked in

various ways to a test gene, such as that encoding β -globin, which was subsequently shown to be efficiently expressed in transfected B cells but not in other cell-types, such as fibroblasts or epithelial cells. However, Schaffner reported that a test gene (the T-antigen gene) linked to a mouse immunoglobulin heavy chain gene enhancer is expressed as efficiently in pre-B and B cells as in plasma cells, even though the level of expression of the endogenous immunoglobulin heavy-chain gene increases dramatically as B cells are stimulated to differentiate into plasma cells.

This observation has led Schaffner and his colleagues to investigate just how the rates of transcription of the immunoglobulin heavy-chain gene compare in pre-B, B and plasma cells. Surprisingly, the rates turn out to be identical at each stage of B-cell differentiation, so the higher level of immunoglobulin heavy chain mRNA in plasma cells than in pre-B and B cells must be the result of some post-transcriptional control mechanism, such as a change in RNA stability.

Further support for this inference comes from the observation reported by M. Neuberger (MRC Laboratory of Molecular Biology, Cambridge) that when the immunoglobulin heavy-chain gene itself is reintroduced into pre-B, B or plasma cells under the control of its own enhancer, its level of expression reflects that of the endogenous gene. Neuberger also reported evidence that the enhancer is not the only component of the immunoglobulin heavy-chain gene to manifest cell-type specificity: the coding region of a β -globin gene linked to the promoter of the immunoglobulin

*The Control of Transcription in Eukaryotes, 10th annual EMBO symposium, 24-27 September 1984, European Molecular Biology Laboratory, Heidelberg, FRG.

heavy-chain gene (which flanks the 5' side of the variable, *V*, region) is properly transcribed only in B cells, even when linked to the SV40 enhancer, which is known to work in many different cell-types. Neuberger suggested that this specificity may be conferred by the conserved sequence that has been found upstream of all immunoglobulin gene *V* regions sequenced to date⁶.

The precise role of enhancer sequences in immunoglobulin gene expression has also been brought into question by two observations. First, the λ light-chain gene does not seem to have one, and second the translocations that frequently have juxtaposed immunoglobulin genes with the *c-myc* gene in B-cell malignancies more often than not have removed the enhancer to the reciprocal translocated chromosome (see ref. 7 and references therein). The implication of the latter observation is that the immunoglobulin gene enhancer cannot be directly responsible for the deregulation of the *c-myc* gene that is thought to have played a key role in these types of cellular transformation⁷.

A similar apparent lack of a requirement for an enhancer is shown by the continued heavy-chain gene expression after spontaneous deletion of regions of the gene that include the enhancer^{8,9}. For one of these cases, that of a δ heavy-chain-producing mouse hybridoma cell line⁹, Schaffner's group has preliminary evidence that the cloned enhancerless gene must be linked to a new enhancer in order to be expressed when it is reintroduced into B cells. On that basis, Schaffner suggests that enhancers are perhaps required for the establishment of transcription, but are not essential for its maintenance. As Schaffner pointed out, this could account for the apparent lack of requirement for an enhancer for the deregulation of translocated *myc* genes.

If this is so, then why are genes that are turned on by steroid hormones turned off again when the hormone is withdrawn, given that the sequences upstream of genes that confer steroid-responsiveness now seem to be just like the inducible enhancers^{10,11}? A possible clue may be in some results reported in Heidelberg by P. Chambon (Institut de Chimie Biologique, Strasbourg).

The principal aim of the experiments reported by Chambon, like those reported by G. Schutz (Institute of Cell and Tumor Biology, Heidelberg) was, however, to answer the puzzle posed in the first sentence of this article — how can one account for the different patterns of expression of steroid-inducible genes in different cell types, all of which contain the appropriate steroid hormone receptors.

For example, the chicken ovalbumin gene is induced by oestrogen in oviduct cells, in which the vitellogenin gene is permanently inactive, whereas in liver cells the ovalbumin gene is permanently inactive and the vitellogenin gene is oestrogen-inducible. Perhaps even more puzzling is

the chicken lysozyme gene, which can be induced by oestrogen, progesterone, testosterone or glucocorticoids in oviduct cells, but is permanently inactive in fibroblasts (which contain glucocorticoid receptors) and is expressed constitutively, albeit at a low level, in macrophages.

From an analysis of the expression in cultured cells of microinjected hybrid genes comprising the SV40 T-antigen coding region linked to flanking sequences of the chicken ovalbumin genes (Chambon) or lysozyme genes (Schutz), it seems that, in addition to hormone-responsive sequences, the genes are controlled by other cell-type specific sequences that further restrict their expression. However, in view of the observation of Chambon's group that the ovalbumin flanking sequences allow constitutive gene expression in injected liver cells, the inactivity of the endogenous gene in liver cells remains a puzzle. Chambon suggests the answer may lie in the establishment of domains of permanently inactive and potentially active chromatin during development.

Chambon also reported that hybrid genes containing up to 295 base pairs of the upstream flanking sequences of the ovalbumin gene are constitutively expressed in oviduct and liver cells, but that if 420 base pairs of flanking sequence are used, expression is blocked in oviduct, but not in liver cells. This block can be overcome by steroid hormones or by including an SV40 enhancer element in the microinjected construct. On the basis of these data, Chambon postulates that oviduct cells contain a repressor that recognizes the upstream 'blocking sequence' and keeps the level of expression of the ovalbumin gene minimally low in the absence of hormone. Could it be that such a repressor also has the role of re-establishing a requirement for enhancer-induction following hormone withdrawal?

The notion that enhancers may be required only for the establishment, and not for the maintenance, of a state of transcription is, however, somewhat difficult to reconcile with a model put forward by K. Yamamoto (University of California, San Francisco) to account for some recent results obtained by his group and which may also provide an explanation of some of the remaining puzzles of the regulation of steroid-inducible genes. In essence, Yamamoto reported that pairs of different enhancer sequences in combination can have synergistic effects on the genes to which they are linked, and that these effects depend on the precise nature of the promoter from which transcription of the gene is initiated.

Yamamoto suggests that these observations can be explained if different enhancers stimulate the supply to the promoter of different specific transcription factors, the synergism resulting from a change in the factor that is rate-limiting for transcription in the presence of one or

other, or both, the different types of enhancer. Following this line of argument, then, the failure of steroid hormones to increase the low constitutive level of lysozyme gene expression in macrophages could be due to the absence of some other factor which is necessary, though not sufficient, to enhance transcription.

This is all speculation, however, and in truth we still have little idea how even steroid hormone-induced enhancers activate transcription. Even the role of the steroid hormone is still obscure for, as M. Beato (Institut für Physiologische Chemie, Marburg) and W. Schrader (Texas Medical Center, Texas) explained, although the presence of hormone is required for the stimulation of transcription, the hormone-receptors bind quite happily to their specific sequences in the absence of hormone. Hopes that some of this ignorance might soon be dispelled were raised in Heidelberg by Yamamoto's news that his group has cloned cDNAs encoding the rat glucocorticoid receptor. □

1. Banerji, J., Olson, L., & Schaffner, W. *Cell* **33**, 729 (1983).
2. Gillies, S.D., Morrison, S.L., Oi, V.T. & Tonegawa, S. *Cell* **33**, 719 (1983).
3. Neuberger, M.S. *EMBO J.* **2**, 1373 (1983).
4. Yaniv, M. *Nature* **297**, 17 (1982).
5. Boss, M.A. *Nature* **303**, 281 (1983).
6. Falkner, F.G. & Zachau, H.G. *Nature* **310**, 71 (1984).
7. Robertson, M. *Nature News and Views* **309**, 585 (1984).
8. Wabl, M.R. & Burrows, P.D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2452 (1984).
9. Klein, S., Sablitzky, F. & Radbruch, A. *EMBO J.* **3**, 2473 (1984).
10. Chandler, V.L., Maler, B.A. & Yamamoto, K.R. *Cell* **33**, 489 (1983).
11. Majors, J.E. & Varmus, H.E. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5866 (1983).

Geoffrey North is an assistant editor of *Nature*.



100 years ago BACTERIOLOGY

AMONG the most striking of the recent rapid advances of science is the development of what we may term bacteriology. While some of this work has been done in this country, by far the greater part has been done abroad, more especially in Germany and France, where its importance is recognised, and where special facilities are afforded by the Governments and various public bodies. In Germany especially, besides the laboratory, supported by the Government, in which Dr Koch works, a number of similar institutions are being established throughout the country; and in France the laboratories of Pasteur and others are established and supported by the Government and by various municipal authorities, every facility for carrying on these researches, and the necessary funds, being provided. In this country, on the other hand, there is no laboratory of the kind, and what work has been done has been by individual investigators working at their own expense and often without suitable accommodation.

From *Nature* **31**, 49, 20 November 1884.

Carbon cycle

Surprise from the shallows

from M. Whitfield

A RARE new insight into the workings of the carbon dioxide cycle is offered by the paper of Byrne *et al.* on page 321 of this issue and its companion, from the same group, which is published in this week's *Science* (226, 1074; 1984). In suggesting, from observations in the North Pacific, that the flux of calcium carbonate shells from the surface layers into the deep oceans has been grossly underestimated, these papers imply a significant shift in our understanding of the oceanic carbon dioxide system.

From an analysis of the particles that settled in large conical traps, deployed at various depths for only a day or so, Byrne and his colleagues were able to show for the first time that pteropod shells, which are made of aragonite — the most soluble form of oceanic calcium carbonate — contribute a flux of calcium carbonate to the deep ocean which is at least as great as that provided by the calcite shells of the foraminifera and coccolithophorids. Furthermore they show that, unlike calcite shells which fall into the sediment and then redissolve only slowly, most of the pteropod shells dissolve in the water column. Earlier measurements had been inconclusive; tow-net sampling had allowed no estimate to be made of the settling rate of pteropods, and long-term sediment traps had inadequately registered the presence of pteropods because their shells had dissolved in the weeks or months before the traps were recovered.

As the shells settle out of the surface layers into deeper water, they resist progressively more corrosive conditions until reaching their lysocline, where they begin to dissolve. Typical lysocline depths for the North Atlantic are 3.2 km for aragonite and 4.6 km for calcite. For the older and more acidic deep waters of the North Pacific, the lysocline depths are typically 0.6 km for aragonite and 0.8 km for calcite. Although calcium carbonate dissolves at and below the lysocline it can still be found in sediments down to the compensation depth, where rates of settling and dissolution just balance.

For each mole of calcium carbonate dissolved, one mole of carbon dioxide is converted into bicarbonate (resulting in an increase in oceanic alkalinity). The new findings make it possible to establish a realistic alkalinity balance for the oceans. Moreover they suggest that the oceans might act as a more rapid and effective sink for fossil-fuel carbon dioxide than had been thought possible when attention was focussed on the dissolution of calcite from the ocean sediments. In that case the response time is measured in centuries or millenia because of the slow circulation of

the carbon dioxide-rich waters down to the compensation depth. To dissolve a significant amount of calcium carbonate, these acidic waters must penetrate the sediments, still further slowing down the process.

The new observations indicate that 5–20 per cent of the smaller pteropod shells will have dissolved by the time they have reached a depth of 1,000 m. This dissolution provides a ready explanation for the mid-water alkalinity maximum that is observed in North Pacific waters substantially supersaturated with respect to calcite, which has previously been ascribed to the possible, but rather unlikely, long-distance transport of high alkalinity water from the edges of the ocean basin.

To close the circle, it is also necessary to consider the enhancement of shell dissolution by the increasing oceanic concentration of carbon dioxide that is resulting from the burning of fossil fuel. The new observations make it clear that this carbon dioxide has only to penetrate to the aragonite lysocline for it to have an effect. Model calculations suggest that it has

already reached a depth of 300 m — close to the aragonite lysocline in the North Pacific — as the result of the slow downward mixing of surface waters. If the dissolution of pteropod shells in relatively shallow waters is enhanced in this way, there will be a corresponding reduction in the alkalinity of the deeper water. In turn, this will accelerate the dissolution of the calcite shells and the larger aragonite shells in the sediment. Furthermore, as the falling shells begin to dissolve, their rate of fall will be reduced and hence their dissolution will become progressively more complete.

On a global basis, this process could accommodate up to one fifth of the current flux of fossil-fuel carbon dioxide and would therefore be much more important than the dissolution of shallow-water, magnesium-rich calcites that has previously been considered the best candidate to account for any rapid response. As oceanic carbon-cycle models become sufficiently sophisticated to incorporate biological processes in a realistic fashion, it will be instructive to see to what extent the oceanic uptake of fossil-fuel carbon dioxide will be influenced by the shift in the alkalinity balance implied by the flux of pteropod shells to the deep ocean and by their tendency to dissolve in transit. □

M. Whitfield is at the Marine Biological Association of the UK, Plymouth PL1 2PB, UK.

Cell biology

A spectrum of spectrins

from A. J. Baines

SOME two years ago I alluded in these columns to the discovery of analogues of the red cell membrane skeleton protein, spectrin, in unexpected places¹. Since then they have been looked for and found — at least by the somewhat rudimentary technique of immunological cross-reaction — in almost every kind of cell that the imagination of cell biologists can run to.

The outsider, trying to follow the trail of discovery, is apt to be put off the scent by a multiplicity of terminologies. Not only are several names (tissue spectrin, caldesmon, and fodrin) still in use for the spectrin-like proteins of non-erythroid cells, but we now have multiple names for their subunits too. For instance, the smaller subunit of brain spectrin, because it binds to the membrane-associated retaining protein, ankyrin, and hybridizes with the larger, alpha subunit of either brain or red cell spectrin, is sometimes termed beta by analogy with the erythrocyte beta chain (for example ref. 2). On the other hand, its electrophoretic mobility in SDS differs from that of erythrocyte beta chain and peptide maps of the two are quite distinct, so what seems to be the same peptide turns up elsewhere under the designation gamma (for example, ref. 3). Ultimately, the relationship between these polypeptides will need to be resolved by

direct comparison of sequences, using that of the greater part of the red cell spectrin molecule⁴ as the standard.

So what does spectrin do in different cells? In parallel with its tissue distribution, some indications of the cellular functions in which it may be involved have accrued. Originally it was supposed that spectrin and its associated proteins, which together make up the red cell membrane skeleton, would perform much the same function in, say, a brain cell as they would in a red cell. Thus it was conceived that a plasma membrane needed a supporting cytoskeletal apparatus to prevent its disintegration, although, to be sure, no brain cell has to withstand the rigours endured by the red cell in circulation.

Complicating this attractively simple model has been the discovery that tissue spectrins are able to enter into associations with proteins not present in the mature red cell, as well as being able to bind protein 4.1, actin, ankyrin and calmodulin. They can bundle microtubules⁵ and bind to at least one of the microtubule-associated proteins (tau)⁶, activities which may be reflected in the antigenic cross-reactivity of spectrin with MAP II, another microtubule-associated protein⁷. Brain ankyrin also binds to microtubules⁸, although in

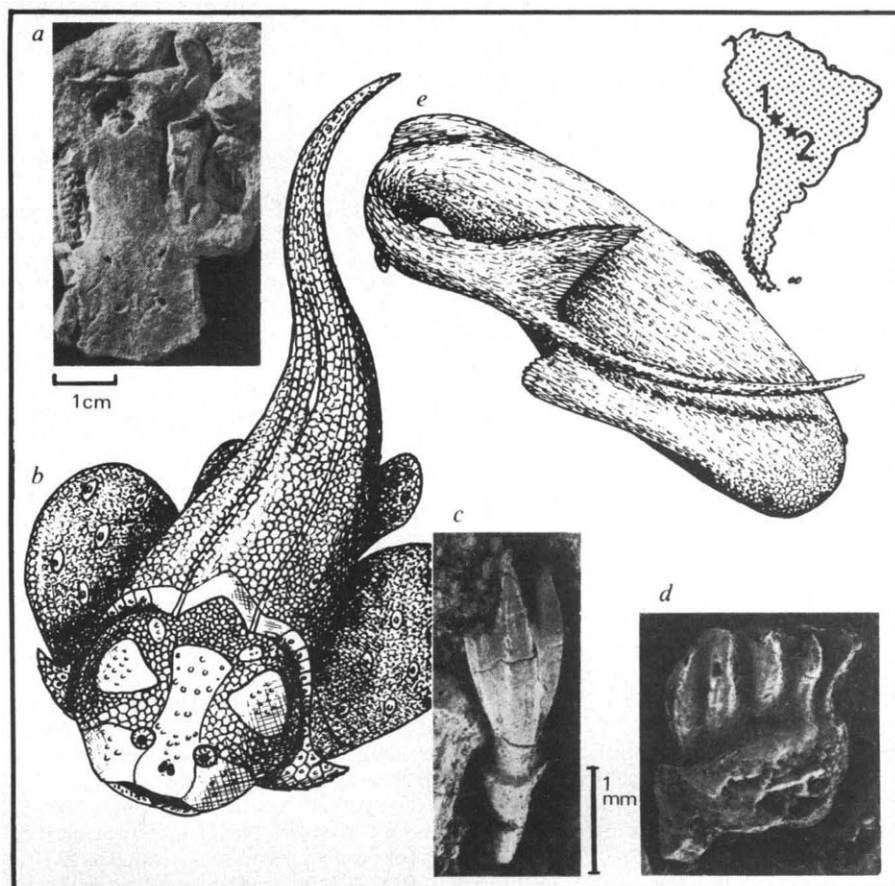
this case an interaction with tubulin dimer is suspected as well: here again there is cross-reaction with another microtubule-associated protein, MAP I⁹. Brain spectrin and ankyrin may indeed be the first examples of membrane-associated microtubule-binding proteins. An indication of an interaction between spectrin and intermediate filaments has come from evidence that when anti-spectrin antibodies are microinjected into cultured cells, not only is spectrin precipitated, but the organization of the intermediate filaments is disrupted¹⁰.

Support for the idea that spectrin is somehow involved in maintaining the physical state of membranes comes from the elegant immunofluorescence studies of Koenig and Repasky¹¹, who show that isolated axons are enriched in spectrin where they are not enclosed by a myelin sheath, that is at the nodes of Ranvier and the synapses. It therefore seems that spectrin is involved in the generation and maintenance of plasma membrane domains. However, this cannot be the whole story since the same group has discovered an association between spectrin, as well as calmodulin and F actin, and the rapidly moving axonal particles (motile varicosities), which are seen in axons regenerating *in vitro*¹². This suggests that the role of spectrin also encompasses the linkage of organelle membranes with cytoskeletal elements. Aunis and Perin¹³ have added a further twist by suggesting the implication of spectrin in chromaffin granule exocytosis. If spectrin is indeed involved in exocytosis then its concentration at the synapse could be taken to imply a role in neurotransmission. A further recent conjecture is that the cleavage of spectrin on post-synaptic membranes, triggered by depolarization-induced events, may be involved in memory storage¹⁴.

So in answer to the question of the cellular functions of spectrin, a general role is now discernible. Spectrin seems frequently to link cytoskeletal elements to plasma or other (for example exocytic) membranes, or to regulate that linkage. Cells may use the linkage in a variety of ways, including the generation and stabilization of plasma membrane domains, the facilitation of intracellular transport (leading in certain cases to exocytosis) and the regulation of the activities of trans-membrane proteins. Watch this space. □

1. Baines, A.J. *Nature News and Views* **301**, 377 (1983).
2. David, J. & Bennett, V. *J. biol. Chem.* **258**, 7757 (1983).
3. Nelson et al. *J. Cell Biol.* **97**, 1271 (1983).
4. Speicher et al. *J. biol. Chem.* **258**, 14931 (1983).
5. Ishikawa, M. *J. Biochem.* **94**, 1209 (1984).
6. Carlier, M.F. et al. *Biochimie* **66**, 305 (1984).
7. David, J. & Bennett, V. *J. biol. Chem.* **257**, 5816 (1982).
8. David, J. & Bennett, V. *J. biol. Chem.* (in the press).
9. Bennett, V. & David, J. *Proc. natn. Acad. Sci.* **78**, 7550 (1981).
10. Mangeat, P. & Burridge, K. *J. Cell Biol.* **98**, 1363 (1984).
11. Koenig, E. & Repasky, E. *J. Neurosci.* (in the press).
12. Koenig, E. et al. *J. Neurosci.* (in the press).
13. Aunis, D. & Perin, D. *J. Neurochem.* **42**, 1558 (1984).
14. Lynch, G. & Baudry, M. *Science* **224**, 1057 (1984).

A.J. Baines is in the Department of Cell Biology and Anatomy, Johns Hopkins University Medical School, Baltimore, Maryland, 21205, USA.



Devonian vertebrates from South America

DEVONIAN vertebrates have been recorded from every continent except South America, where only a few fossils have been tentatively referred to this group (Janvier, P. *Program. intern. Correlat. geol., La Paz* **44** 1; 1978). The most convincing of these is a footprint of an amphibian, from the Devonian of Brazil. Lower and Middle Devonian sites of western and central Bolivia have now yielded scales of agnathans, spines and scales of acanthodians and sharks, and a skull and plates of a placoderm.

The figure illustrates two of the Devonian fishes we have excavated. The Middle Devonian Sicasica Formation, which extends south of La Paz on the Altiplano (site 1 on the map) yielded a well preserved braincase and the jaws of a rhenanid (a), a rare group of placoderms (armoured fishes) known from the Devonian of Europe and North America. There are close similarities between these fossils and a complete rhenanid, *Gemuendina stuerzii* (b) excavated from the Lower Devonian of Hunsrück, FRG. South-east of the Altiplano in the Chuquisaca district (site 2 on the map), the Lower Devonian Catavi Formation crops out along the Rio Grande at Seripona. There, a rich bone-bed has yielded remains of acanthodians, sharks and thelodonts, a group of jawless vertebrates. Two types of thelodont scales (c,d) from this locality may have belonged to a form resembling *Turinina* (e), known from the Lower Devonian of Scotland. The

scales also resemble those of a thelodont recently described from the Devonian of Australia, *Australolepis seddoni*. Similarly, peculiar shark spines from the locality recall those of *Antarctilamna*, from the Middle Devonian of Antarctica and Australia.

Moreover, large acanthodian spines from Seripona, as well as from several other Lower Devonian localities in Bolivia, display the same superficial ornamentation, with alternating broad and narrow ridges without nodes, as the spines of *Sinacanthus*, a large climatiid acanthodian from the Middle Devonian of south China.

Any large-scale biogeographic conclusions would be premature. But at a time when the early history of the Pacific is being vigorously discussed by geophysicists and palaeobiogeographers (Nur, A. & Ben-Avraham, Z. in *Vicariance Biogeography, a Critique*, 341 (Columbia University Press, 1983), the Devonian fishes of South America should provide a valuable tool for estimating Palaeozoic trans-Pacific relationships, particularly since most of them lived very near to the continental shore.

Daniel Goujet and Philippe Janvier, Centre National pour la Recherche Scientifique Institut de Paléontologie, 8 rue Buffon, 75005 Paris, France; M. Suarez-Riglos, Centro de Tecnología Petrolera, Yacimientos Petrolíferos Fiscales Bolivianos, Casilla 7272, Santa-Cruz, Bolivia.

Cell activation

The 'basic' connection

from Walter F. Boron

CHARACTERIZATION of a new biological process almost always includes a description of how it is influenced by changes in pH. In the case of the processes controlled by the enzyme protein kinase C, the tables have been turned; current emphasis is placed on how activation of this kinase affects the pH within the cell (pH_i). Any process that alters pH_i is of potential interest inasmuch as a shift in pH_i could act as a signal to modulate a host of pH-sensitive processes. Activation of protein kinase C is of special interest, because it is believed to mediate the intracellular effects of a wide variety of extracellular signals. On page 371 of this issue Moolenaar *et al.* present evidence that activation of protein kinase C stimulates the plasma membrane Na-H exchanger — which is responsible for the exchange of sodium ions for hydrogen ions across the outer membrane of cells — and thereby leads to an elevation of pH_i in human fibroblasts, HeLa cells and mouse neuroblastoma cells¹. Grinstein *et al.* have reached similar conclusions in experiments on rat thymic lymphocytes². The implication of these data is that the intracellular alkalization induced by the activation of protein kinase C is an important step in the process by which certain extracellular signals elicit their physiological effects.

The Na-H exchanger ultimately affected by protein kinase C is normally involved in the active regulation of pH_i , a housekeeping function performed by all studied animal cells except the erythrocyte. The fundamental acid-base problem faced by these cells is their chronic tendency towards intracellular acidosis, usually caused by the passive influx of acids, such as H^+ , or efflux of bases, such as HCO_3^- , but sometimes caused by the metabolic generation of acid. Uncontrolled falls of pH_i are prevented by ion-transport systems that have the effect of extruding acid from the cell. The invertebrate system appears to exchange external Na^+ and HCO_3^- for internal Cl^- and H^+ , whereas most vertebrate cells use a transporter that exchanges external Na^+ for internal H^+ . Both acid-extrusion mechanisms are ideally suited to their role by being virtually inactive at high pH_i values, and progressively stimulated as pH_i falls below a certain 'threshold'³. This is at least in part due to an intracellular H^+ -binding modifier site, distinct from the site that binds H^+ for transport^{4,5}. The first evidence for modulation of a pH_i regulator was that cyclic AMP stimulates the $Na^+/HCO_3^- - Cl^-/H^+$ exchanger of barnacle muscle⁴. Part of the action of parathyroid hormone on the renal proximal tubule may be due to the inhibition of Na-H exchange by cyclic AMP⁵.

A possible role for Na-H exchange in the

regulation of cell growth by extracellular signals was suggested by experiments in which exposure of quiescent Swiss 3T3 cells to a combination of platelet-derived growth factor (PDGF), vasopressin and insulin, caused a Na^+ -dependent rise in pH_i of about 0.15, as well as a marked stimulation of Na^+ uptake⁷. Subsequently, epidermal growth factor (EGF) was shown to increase pH_i by 0.1–0.2 within ten minutes, but not in the presence of amiloride, an inhibitor of Na-H exchange^{8,9}. Activation of the quiescent cells culminates in cell division several hours later. The notion that an early rise of pH_i plays a role in triggering the cell into division is supported by the observation that the degree of intracellular alkalization induced by growth factors closely parallels the subsequent stimulation of DNA synthesis¹⁰. Thus, the Na-H exchanger, via its effect on pH_i , may be a critical regulator of cell growth.

The cascade of events by which growth factors stimulate Na-H exchange may be as follows: a growth factor, such as EGF¹¹ or PDGF¹², binds at the plasma membrane to a specific receptor (that also has the activity of phosphorylating proteins at tyrosyl residues). A consequence of the binding is the release of Ca^{2+} and increased breakdown of inositol phospholipids. The latter results in the production of diacylglycerol, a potent activator of protein kinase C¹³, which phosphorylates proteins, including the receptor for EGF (see ref. 14), at seryl and threonyl residues. Finally, as indicated by the new data of Moolenaar *et al.*¹, protein kinase C somehow stimulates the Na-H exchanger.

Certain phorbol ester tumour promoters that bear a structural similarity to diacylglycerol are commonly used to stimulate protein kinase C in laboratory investigations. Moolenaar *et al.* show that one such phorbol ester, 12-O-tetradecanoylphorbol-13-acetate (TPA), but not a biologically inactive analogue, causes pH_i to increase by about 0.15 within ten minutes, much as it changes after direct stimulation of fibroblasts by polypeptide growth factors. The TPA-induced increase of pH_i is Na^+ dependent and is blocked by a potent amiloride analogue, indicating that it is caused by stimulation of a Na-H exchanger. Furthermore, TPA does not induce a pH_i increase in fibroblasts that have already been stimulated by PDGF. Inasmuch as TPA bypasses the growth-factor receptor/kinase step of the proposed cascade, a simplistic view would be that stimulation of the Na-H exchanger does not require the release of either Ca^{2+} or diacylglycerol *per se*, but only the activation of protein kinase C. Indeed, the

observations of Moolenaar *et al.* on Na-H exchange, as well as those of others on secretory processes (see ref. 15), indicate that TPA can produce its pharmacological effect without a measurable rise of intracellular Ca^{2+} .

The simplest explanation for the effects of TPA on pH_i , and the one proposed by Moolenaar *et al.*, is that protein kinase C directly phosphorylates the Na-H exchanger. This remains to be established. The kinetic effect of TPA on the Na-H exchanger could be to increase its pH_i threshold — that is to increase the value to which pH_i must rise before the Na-H exchanger shuts off — and/or to increase its sensitivity to acidosis once pH_i falls below the threshold. Moolenaar *et al.* favour the former. This is an attractive possibility, although it is easy to imagine how phosphorylation near the intracellular modifier site could alter either the pH_i threshold or the sensitivity of the Na-H exchanger.

Finally, it is intriguing to speculate on the possible physiological role of the stimulated Na-H exchanger. Most of the experiments on the effects of growth factors on pH_i have been conducted in HCO_3^- -free media. Elsewhere, Moolenaar *et al.* have reported that unstimulated human fibroblasts do not have a detectable HCO_3^- transport system¹⁶. However, L'Allemain *et al.*¹⁰ report that, whereas a functional Na-H exchanger is necessary for stimulating the mitogenesis of Chinese hamster lung fibroblasts that are incubated in HCO_3^- -free media, it is not necessary when the cells are in HCO_3^- -containing media. Is the Na-H exchanger in these hamster cells functionally replaced by a HCO_3^- transport system? Other, more general questions arise. Is a sudden increase of pH_i really necessary to trigger cell growth? Or do growth factors stimulate Na-H exchange for other reasons, such as to enhance the housekeeping ability of the cell to regulate pH_i in anticipation of the increased requirements of the cell on metabolic activation? □

1. Moolenaar, W.H., Tertoolen, L.G.J. & de Laat, S.W. *Nature* **312**, 371 (1984).
2. Grinstein, S., Cohen, S., Goetz, J.D., Rothstein, A. & Gelfand, E. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
3. Boron, W.F. *J. membrane Biol.* **72**, 1 (1983).
4. Boron, W.F., Russell, J.M., Brodwick, M.S., Keifer, D.W. & Roos, A. *Nature* **276**, 511 (1978).
5. Aronson, P., Nee, J. & Suhm, M. *Nature* **299**, 161 (1982).
6. Kahn, A.M., Dolson, G.M., Bennett, S.C. & Weinman, E.J. *Kidney Internat.* **25**, 289 (1984).
7. Schuldiner, S. & Rozengurt, E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7778 (1982).
8. Moolenaar, W.H., Tsien, R.Y., van der Saag, P.T. & de Laat, S.W. *Nature* **304**, 645 (1983).
9. Rothenberg, P., Glaser, L., Schlesinger, P. & Cassel, D. *J. biol. Chem.* **258**, 12644 (1983).
10. L'Allemain, G., Paris, S. & Pouyssegur, J. *J. biol. Chem.* **259**, 5809 (1984).
11. Cohen, S., Carpenter, G. & King, L., Jr. *J. biol. Chem.* **255**, 4834 (1980).
12. Heldin, C.-H., Ek, B. & Ronnstrand, L. *J. biol. Chem.* **258**, 10054 (1983).
13. Nishizuka, Y. *Nature* **308**, 693 (1984).
14. Hunter, T. *Nature News and Views*, **311**, 414 (1984).
15. Baker, P.F. *Nature News and Views* **310**, 629 (1984).
16. Moolenaar, W.H., Tertoolen, L.G.J. & de Laat, S.W. *J. biol. Chem.* **259**, 7563 (1984).

Walter F. Boron is in the Department of Physiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA.

Territory, size and weight gain

SIR — A recent paper by Carpenter *et al.*¹, on territory size and weight gain, has been received with enthusiasm by Diamond². One must share his attitude towards their elegant experimental technique for weighing birds undisturbed in the wild. However, I believe more caution should be exercised with regard to their results and interpretation. Diamond quotes the example of bird WP whose rate of weight gain was measured at three territory sizes and found to be greatest at the intermediate and final size. "The results with WP and other rufous hummingbirds", Diamond says, "reveal that the birds adjust their territory size until they have maximized the rate of weight gain". But WP was in fact the only individual bird observed to have a higher rate of weight gain for an intermediate territory size and even in this case there is no direct evidence that the rate of weight gain achieved was a maximum. The justification for concluding that the birds did maximize their energy input was that in four out of five cases the highest weight gains achieved were on the last day or days of the birds' stay on the study site. Clearly this only demonstrates a level at which birds no longer attempt to increase their net energy input, not that that level is a maximum. The argument for a general "humped" relationship between territory size and weight gain is that some birds showed increases in weight gain with increasing territory size and some with decreasing territory size — a much weaker case than that implied by using the example of WP alone.

Alternative explanations for the observations are possible. For example, suppose that the rate of weight gain is related to the bird's body weight. The observed increase in rate of gain towards the end of each bird's stay is predicted, but is independent of territory size. This would also explain the unusually high rate of weight gain of bird BGR — the bird with the highest initial body weight.

A further problem arises with the proposed adjustment of territory size by "trial and error". Suppose a bird has a sub-optimal energy gain from its initial territory size. Given a bell-shaped relationship between territory size and weight gain, how does the bird decide in which direction to alter its territory size? Even after one alteration made in a random direction the bird still has no good grounds on which to make its next decision.

Apart from the valuable technical innovation, the most important point which comes out of this work is, surely, that the birds do not behave optimally. On arrival at the study site, they do not know and cannot know what the optimum territory is. It depends on many unknown factors, such as nectar production by the flowers and the number and quality of the birds'

neighbours. They spend at least the first few days, if not all of their stay, in territories that are either too large or too small and the best they can do is to try to increase their efficiency as time goes on.

R. K. BUTLIN

*School of Biological Sciences,
University of East Anglia,
Norwich NR4 7TJ, UK*

1. Carpenter, F. I., Paton, D. & Hixon, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7259 (1983).
2. Diamond, J. *Nature* **311**, 603-604 (1984).

Pigeons, canaries and problem-solving

SIR — I regret that I must question the accuracy of Pastore's¹ recent letter concerning my report of problem-solving behaviour in pigeons². He suggests that I was remiss in not citing his early studies in which a canary was confronted with several Köhler-type problems. He implies that the canary solved both Köhler's one-box problem and Köhler's stacking problems in an "insightful" fashion, but the original reports of this work^{3,4} do not confirm that.

The single canary that Pastore confronted with a variant of the one-box problem required 25 reinforced trials before it could move its box (a cardboard "prism") smoothly to the correct position on the floor of the chamber. So the behaviour that superficially resembled that of Köhler's chimps was, as Pastore originally reported, learned in virtually the same haphazard fashion as the escapes of Thorndike's⁵ cats. Ironically, the performance of the canary was exactly that which Köhler dismissed as mere "trial and error". By no reasonable criteria could the canary's performance be considered "insightful".

On the stacking problem, in which stacking behaviour is also established haphazardly after many reinforced trials, Pastore³ originally reported that "in the crucial trial, when both prism and box were out of position, the canary seemed to be unable to stack them in a meaningful way. Actually, the canary did stack prism and box appropriately in only 10 of 100 trials" (p. 289). No mention is made of success in stacking a larger number of boxes, contrary to Pastore's¹ recent statement. A second canary could not complete even the preliminary stages of the experiment. Praxists and psychologists long ago passed judgment on these modest demonstrations: they are almost universally uncited in the literature relevant to my research.

In contrast, my colleagues and I reported a systematic study with 11 pigeons in which all 3 with relevant training histories solved the box-and-banana problem in a dramatic human-like fashion the first time they were confronted with it. The performances satisfied all of the traditional criteria of "insight": periods of apparent confusion were followed by sudden, rapid, and entirely appropriate performances. By

systematically varying the training histories of other birds, we also determined the possible contributions that a variety of different experiences had made to success in the problem. Finally, we offered a running account of the novel performances in terms of empirically validated principles.

ROBERT EPSTEIN

*Cambridge Center for
Behavioural Studies,
11 Ware Street, Cambridge,
Massachusetts 02138, USA*

1. Pastore, N. *Nature* **310**, 18 (1984).
2. Epstein, R., Kirshnit, C., Lanza, R. P. & Rubin, L. C. *Nature* **308**, 61-62 (1984).
3. Pastore, N. *J. comp. physiol. Psychol.* **47**, 288-289 (1954).
4. Pastore, N. *Scient. Am.* **192**, 72-79 (1955).
5. Thorndike, E. L. *Psychol. Rev. Monogr.* **2**, No. 4 (1898).

DNA sequence selection by eye

SIR — The recent correspondence concerning various ways of representing DNA sequences^{1,2} prompts me to describe a simple method for finding sequence overlaps by eye. Using the 4,361-nucleotide sequence of the plasmid pBR322 as a test system, it is possible with a little practice to locate a randomly selected sequence of about 100 nucleotides in less than 2 minutes.

The method requires four ballpoint pens of different colours and the sequence (or collection of sequences) with which the test sequence is to be compared is coded as follows. Every time the sequence AG (an arbitrary choice) occurs, the next base is coloured, the colour used depending on the next base again (G green, A blue, C black, and T red). Thus AGCT appears as a red C, AGGG appears as a green G and so forth. This process takes about 10 minutes per kilobase, but it only has to be done once for each new sequence determined.

To locate the test sequence, an AG is selected at random and about ten adjacent downstream bases are written with a sharp pencil on the edge of a piece of paper. The appropriate base adjacent to the G is coloured and the sequence is checked against each occurrence of that coloured letter, on average four times per kilobase.

If two AG sequences are close together on the test sequence, the pattern of two coloured letters separated by a certain number of bases makes searching faster and also highly specific. Such a pattern would only occur at random three times in the entire Epstein-Barr virus sequence.

If the sequence is not found, the complementary strand is searched by locating one or more CT sequences, writing down the ten or so downstream bases on the complementary strand and proceeding as before.

M. G. BURDON

*Department of Biochemistry
and Microbiology,
University of St Andrews,
St Andrews, Fife KY16 9AL, UK*

1. Hayashi, K. & Munakata, N. *Nature* **310**, 96 (1984).
2. Lathe, R. & Findlay, R. *Nature* **311**, 610 (1984).

Transformation by the p21^{ras} protein

SIR — Ralph¹ has questioned the validity of certain speculative comments of mine concerning the mechanisms by which p21^{ras} proteins transform mammalian cells. He has obviously misunderstood the point of my remarks.

In my article, I drew an analogy between new findings showing loss of GTPase activity in the transforming protein and the fixation of adenylate cyclase Gs protein in the GTP state after cholera toxin binding. In case it was not made clear in the article, perhaps I should state that there is no evidence of a functional link between the adenylate cyclase system and p21^{ras}, and therefore for a direct role of cyclic AMP regulation in *ras*-mediated transformation. Examples of some of the known effects of cholera toxin on cell physiology and proliferation were mentioned solely to illustrate the profound consequences which result from fixation of the adenylate cyclase G-protein in the GTP form. As stated in my article, which (if any) specific effector systems involve p21^{ras} in signal transduction remains to be established.

R. F. NEWBOLD

Chester Beatty Laboratories,
London SW3 6JB, UK

1. Ralph, R. K. *Nature* 311, 515 (1984).
2. Newbold, R. F. *Nature* 310 628-629 (1984).

The gyroscope test of relativity

SIR — Your account of the "relativity gyroscope experiment"¹ does not go beyond that of Schiff², who proposed that two effects, the geodetic (de Sitter or spin-orbit) and motional (Lense-Thirring or spin-spin) precessions, could be measured separately by the use of at least two gyros in free fall in a perfect polar orbit. An accuracy of 0.3 milliarc-seconds per year is the desired goal. However, there are many other contributions which are well above 0.3 marc s yr⁻¹ (for a review see ref. 3).

The goal of the experimentalists is to make a sphere of fused quartz accurate to one part in 10⁷, but even a perfect sphere is distorted when set spinning, acquiring a quadrupole moment, which for an altitude of 500 miles, will contribute more than 0.3 marc s yr⁻¹ to the gyro drift rate if the spin axis is more than 20 minutes of arc away from either the orbit plane or the perpendicular to the orbit plane⁴.

Second, the Earth's quadrupole moment will contribute an amount of 4 marc s yr⁻¹, more than ten times the desired accuracy, and, more subtly, distorts the satellite orbit so that, in averaging over a polar orbit, one obtains an additional "indirect contribution" of 1.33 marc s yr⁻¹.

Third, the contribution due to the Sun will give rise to a geodetic precession of the gyro (due to interaction of the spin of the gyro with its orbital angular momentum about the Sun), with the relatively

large value of 19.2 marc s yr⁻¹, while the deflection of light from the reference star (Rigel in Orion) can cause an apparent drift of the gyroscope of up to 14.4 marc s yr⁻¹.

The relativity gyro experiment test of general relativity is thus more complex than originally envisaged. Hence, there is a need for other tests of the geodetic and motional precessions, particularly the latter. A promising suggestion in this direction is the proposal of Scully and co-workers^{5,6} to use a ring laser interferometer to test both the geodetic and the motional precessions, which has the attraction that it can be carried out on Earth, while another proposal is to use a Foucault pendulum at the South Pole⁷. Also, the binary pulsar PSR1913+16 may provide a test of the spin-orbit precession and our calculations⁸ predict a precession rate for the pulsar spin axis of 1.23° yr⁻¹, based on the data of Taylor and Weisberg⁹.

B. M. BARKER

Department of Physics
and Astronomy,
The University of Alabama
University, Alabama 35486, USA

R. F. O'CONNELL

Department of Physics
and Astronomy,
Louisiana State University,
Louisiana 70803, USA

1. Maddox, J. *Nature* 310, 363 (1984).
2. Schiff, L. I. *Proc. natn. Acad. Sci. U.S.A.* 46, 871 (1960); *Phys. Today* 14, 42 (1961).
3. Barker, B. M. & O'Connell, R. F. *Gen. Rel. Grav.* 11, 149 (1979).
4. Barker, B. M. & O'Connell, R. F. *Phys. Rev. D* 11, 711 (1975).
5. Scully, M. O., Zubairy, M. S. & Haugan, M. P. *Phys. Rev. A* 24, 2009 (1981).
6. Schleich, W. & Scully, M. O. in *Les Houches lectures* (in the press).
7. Braginsky, V. B., Polnarev, A. G. & Thorne, K. S. *Phys. Rev. Lett.* 53, 863 (1984).
8. Barker, B. M. & O'Connell, R. F. *Astrophys. J. Lett.* 199, L25 (1975); *Phys. Rev. D* 12, 329 (1975).
9. Taylor, J. H. & Weisberg, J. M. *Astrophys. J.* 253, 908 (1982).

Mouse Ig κ sequence elements in *Drosophila*

SIR — Two sequence elements may be needed for correct transcription of the mouse immunoglobulin κ gene. They are: TNATTTGCAT (dc), its inverted and complementary ATGCAAATNA (cd), upstream of the H-chain gene and TGCA^cCTGTGNCCAG (pd)¹. The same sequences (8 out of 10 base pairs) ATTTGCAT and ATGCAAAT have been found by Parslow *et al.* independently². Sequences related to the above elements were found upstream of all human and mouse κ and λ variable region genes and within the mouse heavy-chain enhancer in a homology of 50–100%. The same elements are present in high homology in chicken ovalbumin gene, sea urchin histone gene cluster and in other genes¹. I report here that the elements dc and cd are present upstream of the homoeotic gene *ftz*³ of *Drosophila*.

The element AGGCAAATAC is found upstream the coding region of *ftz* (first

nucleotide position -794)³. This has 78% homology with the cd element ATGCAAATNA. The sequence ATGTTTGCAT is also found upstream of the coding region of *ftz* (first nucleotide -854)³ and shares 78% homology with the dc element TNATTTGCAT. Other sequence elements can also be found upstream or within the *ftz* gene with less homology with the dc, cd or pd 50–67%, but this does not seem to be significant since dc, cd or pd can be found in many genes when homology less than 70% is considered (F. G. Falkner and H. G. Zachau, personal communication).

The dc and cd elements are highly conserved among many genes. Their location within the immunoglobulin genes suggests functional relation¹. The homoeotic gene *ftz* is associated with *Drosophila* development and contains a highly conserved region, the homoeo box, which has been found in other homoeotic genes — *Antp* and *Urb* of *Drosophila* as well as in a wide spectrum of animals. The relationship of conserved sequences needed for transcription to upstream sequences needed for segmentation during development is unknown. It is, of course, possible that the homologies might be chance events, and more information is needed before evolutionary or functional relationships are considered.

La Jolla Cancer Research Foundation,
Cancer Research Center,
La Jolla, California 92037, USA

1. Faulkner, F. G. & Zachau, H. G. *Nature* 310, 71-74 (1984).
2. Parslow, T. G. *et al. Proc. natn. Acad. Sci. U.S.A.* 81, 2650-2654 (1984).
3. Laughton, A. & Scott, M. P. *Nature* 310, 25-31 (1984).
4. Shepherd, J. C. W. *et al. Nature* 310, 70-71 (1984).
5. McGinnis, W. *et al. Cell* 37, 403-408 (1984).
6. Carrasco, A. E. *et al. Cell* 37, 409-414 (1984).

FAULKNER AND ZACHAU REPLY — We agree that the occurrence of dc-related elements upstream of genes deserves attention. We quote some examples for that in our previous paper¹ and others in a forthcoming one (F. G. F., E. Neumann and H. G. Z., manuscript in preparation). At this point we would like to mention that a sequence fully homologous to dc was found upstream of all or most histone H2B genes², a point which we had missed previously. In general, we searched sequence libraries for dc at the 89% homology level since at lower levels too many sequences are picked up. But also dc (and pd) related sequences with lesser homology may be interesting if they can be correlated with putative regulatory functions. One example of this which we recently encountered is the finding of the dc and cd related sequences in the regulatory regions upstream of catabolite-sensitive genes of *Escherichia coli*^{3,4}.

1. Falkner, F. G. & Zachau, H. G. *Nature* 310, 71-74 (1984).
2. Harvey, R. P., Robins, A. J. & Wells, J. R. E. *Nucleic Acids Res.* 10, 7851-7863 (1983).
3. de Crombrughe, B., Busby, S. & Bruc, H. in *Biological Regulation and Development* Vol 3B (eds Goldberger, R. F. & Yamamoto, K. R.) 129-167 (1984).

Inositol trisphosphate, a novel second messenger in cellular signal transduction

Michael J. Berridge* & Robin F. Irvine†

* AFRC Unit of Insect Neurophysiology and Pharmacology, Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

† Department of Biochemistry, AFRC Institute of Animal Physiology, Babraham, Cambridge CB2 4AT, UK

There has recently been rapid progress in understanding receptors that generate intracellular signals from inositol lipids. One of these lipids, phosphatidylinositol 4,5-bisphosphate, is hydrolysed to diacylglycerol and inositol trisphosphate as part of a signal transduction mechanism for controlling a variety of cellular processes including secretion, metabolism, phototransduction and cell proliferation. Diacylglycerol operates within the plane of the membrane to activate protein kinase C, whereas inositol trisphosphate is released into the cytoplasm to function as a second messenger for mobilizing intracellular calcium.

THE behaviour of all cells from one instant to another is governed by signalling systems that translate external information into a limited repertoire of internal signals, second messengers. Receptors on the cell surface function as molecular antennae, detecting external information (for example, hormones, growth factors, neurotransmitters or light) which is then transduced and amplified into the second messengers which control many cellular processes such as metabolism, secretion, contraction, phototransduction and cell growth. Interest in this field of signal transduction is growing for several reasons, but perhaps not least because an imbalance of second messengers may be responsible for normal cells becoming cancerous.

The signalling system using cyclic AMP as the second messenger may be the best known. Until recently, we knew much less about the nature of the second messengers used by another major signalling pathway that utilizes the inositol lipids as part of a transduction mechanism. Stimulated inositide metabolism was discovered by Hokin and Hokin¹, who showed that the incorporation of ³²P into phospholipids in pancreas was stimulated by acetylcholine. It later became clear that a stimulated catabolism of inositol lipids occurs in many different cells in response to a wide range of external signals²⁻⁴. Durell and co-workers⁵ suggested that this stimulated metabolism may be a part of receptor function, while Michell subsequently proposed that inositide metabolism is concerned with the mode of action of those agonists which raise intracellular calcium². Although this hypothesis is not universally accepted in its entirety^{6,7}, the many coincidences between calcium-mobilizing agonists and stimulated inositide turnover⁸, and many negative 'controls' (that is, receptors which do not mobilize calcium do not stimulate inositide catabolism) has led to a widespread acknowledgement that calcium and inositides are closely linked. It has also become clear that the mobilization of calcium from intracellular stores is an important source of activator calcium following stimulation of many different cells, including oocytes⁹, Swiss 3T3 cells¹⁰, pancreas¹¹, parotid¹², GH₃ pituitary cells¹³, blood platelets¹⁴, insulin-secreting β -cells¹⁵, liver¹⁶ and smooth muscle¹⁷. Thus, there has been a missing link not only between calcium and inositides, but also between the plasma membrane (where the relevant receptors lie) and the internal stores of calcium. Our hypothesis is that inositol trisphosphate (InsP₃) is that missing link.

Formation of inositol trisphosphate

The major metabolic pathways responsible for the formation and degradation of InsP₃ are illustrated in Figs 1 and 2 (for reviews see refs 3, 18). Phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P₂) is one of the inositol lipids located in the inner leaflet of the plasma membrane (Fig. 1), and is formed by a

two-stage phosphorylation of phosphatidylinositol, which is first phosphorylated at the 4-position of its inositol head group by a specific kinase to form phosphatidylinositol 4-phosphate (PtdIns(4)P); this is in turn further phosphorylated at the 5-position to give PtdIns(4,5)P₂, which is the immediate precursor used by the receptor mechanisms. As well as the two kinases responsible for the stepwise phosphorylation of phosphatidylinositol to PtdIns(4,5)P₂ (Fig. 1), there are corresponding phosphomonoesterases which convert PtdIns(4,5)P₂ back to phosphatidylinositol (Fig. 2) These kinases and phos-

Table 1 Summary of the tissues which respond to external stimuli by a change in the hydrolysis of PtdIns(4,5)P₂

Tissue	Stimulus
Rabbit iris smooth muscle	Acetylcholine ^{20,21*}
Brain	Acetylcholine ^{50*}
Parotid	Acetylcholine ^{50*,105*,106*} , substance P ^{50*}
Pancreas	Acetylcholine ^{72*} , caerulein ^{72*}
Avian salt gland	Acetylcholine ¹²⁰
Liver	Adrenaline ⁷¹ , ATP ⁷¹ , vasopressin ^{22,28*,30,58*} , angiotensin ⁷¹ , platelet activating factor ¹²¹
Blowfly salivary gland	Serotonin ^{23*,24*,122*}
Blood platelets	ADP ¹²³ , thrombin ^{124*,125} , platelet activating factor ¹²⁶
Sympathetic ganglion	Vasopressin ^{127*}
GH ₃ pituitary tumour cells	Thyrotropin-releasing hormone ^{25*,26*,27*,29*}
Adrenal cortex	Angiotensin ⁵⁴
Pancreatic islets	Glucose ¹²⁸
Neutrophils	f-Methionyl-leucyl-phenylalanine ^{129,130}
HL-60 leukaemic cells	f-Methionyl-leucyl-phenylalanine ^{131*}
Leukocytes	Pseudomonas leukocidin ¹³²
T-lymphoblastoid cell	Phytohaemagglutinin ¹³³
Leukaemic basophils	Antigen ^{134*}
Swiss 3T3 cells	PDGF ^{63*,88*} , bombesin ^{88*} , vasopressin ^{88*}
Neuroblastoma-glioma hybrid NG108-15	Bradykinin ¹³⁵
Photoreceptors	Photons ^{42*,43}
Sea urchin eggs	Spermatozoa ⁸⁵

PtdIns(4,5)P₂ hydrolysis was usually measured by following changes in the distribution of ³H-inositol or ³²P.

* Papers describing increases in the formation of labelled inositol trisphosphate (likely to be a mixture of the two isomers¹⁹). For more extensive lists of tissues where inositide turnover in general (rather than PtdIns(4,5)P₂ in particular) is stimulated, see refs 2, 8, 22.

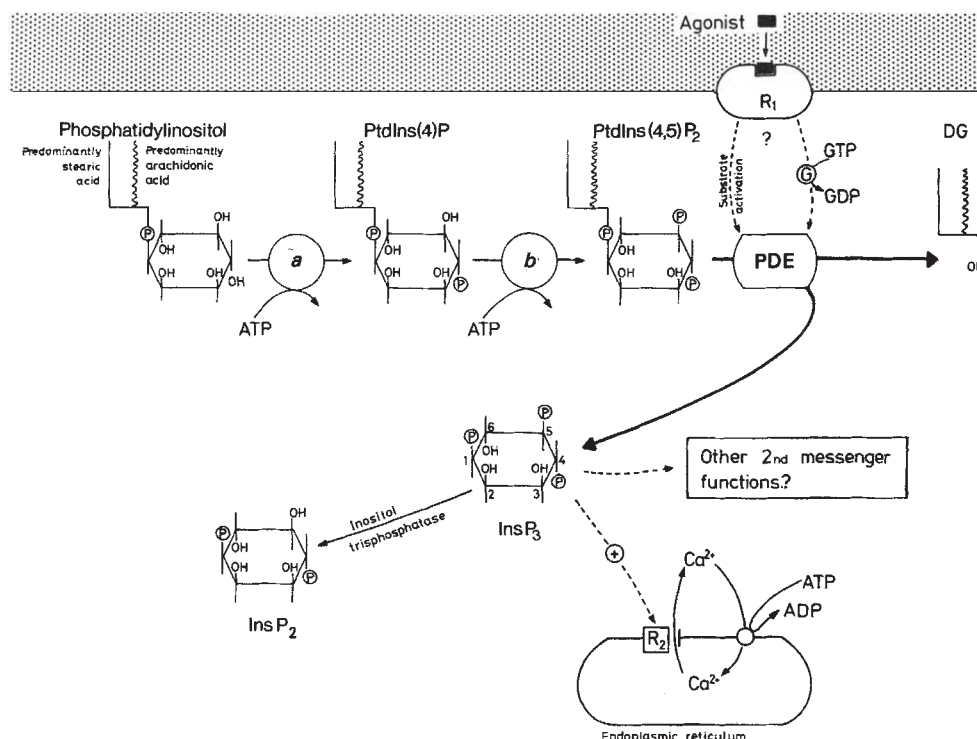


Fig. 1 The proposed role of inositol trisphosphate as an intracellular second messenger. Agonists bind to external receptors (R₁) to stimulate the hydrolysis of PtdIns(4,5)P₂ by a phosphodiesterase (PDE) to form diacylglycerol (DG) and InsP₃. The latter may have a number of second messenger functions one of which is to bind to a specific receptor (R₂) on the endoplasmic reticulum to release calcium. The action of InsP₃ is curtailed by an inositol trisphosphatase which removes a phosphate from the 5-position to form InsP₂. The enzymatic pathways responsible for resynthesising phosphatidylinositol from diacylglycerol and InsP₃ are shown in Fig. 2. *a*, Phosphatidylinositol kinase; *b*, PtdIns(4)P kinase; PDE, PtdIns(4,5)P₂ phosphodiesterase.

phomonoesterases constitute two linked phosphatidylinositol/PtdIns(4,5)P₂ futile cycles whereby phosphates are constantly being added to and removed from the 4- and 5-positions of the inositol head group. Such futile cycles are metabolically expensive, which may explain why the operation of this signal transduction mechanism is rapidly curtailed by metabolic inhibitors⁴. Occupation of a receptor by an agonist diverts PtdIns(4,5)P₂ out of the futile cycle towards a phosphodiesterase (phospholipase C), by which it is cleaved into *sn*-1,2-diacylglycerol and InsP₂ (Figs 1, 2).

In passing, it should be noted that two naturally occurring isomers of inositol trisphosphate have been identified, *D*-*myo*-inositol 1,4,5-trisphosphate and *D*- or *L*-*myo*-inositol 1,3,4-trisphosphate¹⁹. In most of the studies summarized in Table 1, where increases in inositol trisphosphate following cell stimulation have been reported, the measurements probably include both isomers. Because the 1,3,4-isomer has only just been described and its function is unknown, we have concentrated our attention on the 1,4,5-isomer, which we abbreviate as InsP₃.

Stimulated catabolism of one of the polyphosphoinositides was first detected by Durell *et al.*⁵, but other workers later documented in detail the phosphodiesterase cleavage of PtdIns(4,5)P₂ in iris smooth muscle to form InsP₃ (refs 20, 21). At first, however, this reaction was thought to be a consequence of the calcium signal in the stimulated tissue, rather than its cause. A crucial advance which changed our concept of inositide metabolism and its function, was the discovery that, in hepatocytes stimulated by vasopressin, decreases in PtdIns(4,5)P₂ precede changes in phosphatidylinositol and are largely independent of external calcium²². The primary receptor-stimulated event is probably the hydrolysis of PtdIns(4,5)P₂ to yield *sn*-1,2-diacylglycerol (as does phosphodiesterase hydrolysis of any glycerophospholipid) and—uniquely to PtdIns(4,5)P₂—InsP₃ as a water-soluble product^{3,4,22}. There are now many studies which support the notion that this hydrolysis of PtdIns(4,5)P₂ is a common response by many different kinds of cells to a wide variety of external stimuli (Table 1). The demonstration that the formation of InsP₃ precedes that of inositol monophosphate (the expected product of phosphatidylinositol hydrolysis) in *Calliphora* salivary glands^{23,24} and clonal GH₃ pituitary cells^{25–27} provides additional evidence that PtdIns(4,5)P₂ is the primary substrate for the receptor. Most importantly, InsP₃ formation in *Calliphora* salivary glands is

stimulated by serotonin with no apparent lag, whereas the onset of the calcium-dependent physiological response is delayed by at least 1 s. This led to the suggestion that InsP₃ is the second messenger for calcium mobilization^{23,24}. Increases in InsP₃ following cell stimulation also precede physiological responses in liver²⁸ and GH₃ cells²⁹, thus confirming that InsP₃ formation is fast enough for it to function as a second messenger.

One unsolved problem in the formation of InsP₃ and diacylglycerol is that of how receptors are coupled to the PtdIns(4,5)P₂ phosphodiesterase (Fig. 1). A recent report that vasopressin can stimulate the breakdown of PtdIns(4,5)P₂ in isolated membranes³⁰ may help to elucidate the transduction mechanism. Agonists may induce a conformational change in the receptor, which in turn perturbs the membrane sufficiently to make PtdIns(4,5)P₂ accessible to the phosphodiesterase³¹. PtdIns(4,5)P₂ phosphodiesterase may also be controlled by coupling of receptors to the enzyme through a GTP-binding protein as in the mechanism used to link receptors to adenylate cyclase^{32,33}.

The first indications that a GTP-binding protein may function in signal transduction for calcium-mobilizing receptors were reports that GTP reduces the affinity of noradrenaline for α_1 -receptors³⁴ and that of carbachol for muscarinic receptors³⁵. In adrenoreceptors, pertussis toxin reduces the modulatory effect of GTP on α_2 -receptors, which act by inhibiting adenylate cyclase, but has no effect on the calcium-mobilizing α_1 -receptors, suggesting that these two receptor mechanisms are regulated by different GTP-binding proteins³⁶.

A more direct link has emerged from studies on permeabilized blood platelets, where either GTP or its non-hydrolysable analogues GTP γ S or GppNHp enhances the calcium sensitivity of the secretory mechanism³⁷, apparently by stimulating the formation of diacylglycerol³⁸. The stimulatory effect of guanine nucleotides introduced into various intact cells can probably be reinterpreted on the basis that they stimulate the hydrolysis of inositol lipids³⁷. For example, stimulation of mast cell secretion by non-hydrolysable GTP analogues³² may result from stimulation of the normal receptor mechanism³⁷. Likewise, the ability of GTP γ S to excite *Limulus* photoreceptors may also depend on the activation of a GTP-binding protein^{39,40}. Such a protein has been identified in squid photoreceptors⁴¹, where light stimulates the turnover of PtdIns(4)P and PtdIns(4,5)P₂ (refs 42, 43). Finally, the *ras* gene product, p 21, which is a membrane-bound

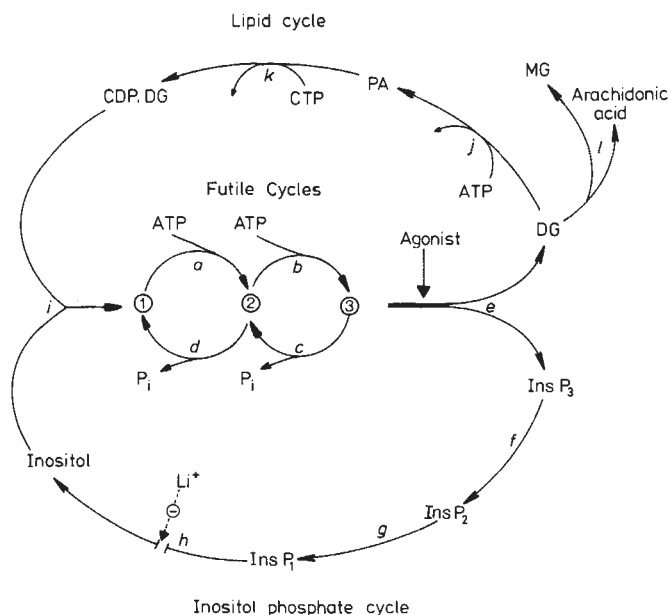


Fig. 2 Inositol lipid metabolism and signal transduction. An inositol phosphate cycle converts InsP_3 to free inositol which interacts with CDP-diacylglycerol (CDP.DG) to re-form phosphatidylinositol (1). The CDP-diacylglycerol is the product of a lipid cycle which channels diacylglycerol back to phosphatidylinositol. Diacylglycerol can also be hydrolysed by a diacylglycerol lipase with the release of arachidonic acid and monoacylglycerol (MG)¹³⁶. The phosphatidylinositol/PtdIns(4,5) P_2 futile cycles maintain a supply of the PtdIns(4,5) P_2 which is hydrolysed as part of the bifurcating receptor mechanism to generate the second messengers diacylglycerol and InsP_3 . a, Phosphatidylinositol kinase; b, PtdIns(4)P kinase; c, PtdIns(4,5) P_2 phosphomonoesterase; d, PtdIns(4)P phosphomonoesterase; e, PtdIns(4,5) P_2 phosphodiesterase; f, InsP_3 phosphatase; g, InsP_2 phosphatase; h, InsP_1 phosphatase; i, CDP-diacylglycerol inositol phosphatidate transferase; j, diacylglycerol kinase; k, CTP phosphatidate cytidyl transferase; l, diacylglycerol lipase. PA, phosphatidic acid. 1, Phosphatidylinositol; 2, PtdIns(4)P; 3, PtdIns(4,5) P_2 .

protein capable of both binding and hydrolysing GTP⁴⁴⁻⁴⁷, may function like the postulated GTP-binding protein to link receptors to PtdIns(4,5) P_2 phosphodiesterase. Consistent with this notion is the finding that epidermal growth factor (EGF) increases the binding of GTP to *ras* p 21⁴⁸, suggesting that the *ras* gene product is capable of interacting with cell-surface receptors. The well established role of GTP-binding proteins in controlling cyclic AMP production may thus be extended to include the regulation of other transduction mechanisms such as InsP_3 formation by these calcium-mobilizing receptors.

Removal of inositol trisphosphate

The intracellular levels of diacylglycerol and InsP_3 are determined by a balance between their rate of formation from PtdIns(4,5) P_2 and their rate of removal by two pathways which channel them back to phosphatidylinositol (Fig. 2). A lipid cycle begins with diacylglycerol kinase (Fig. 2j) converting diacylglycerol to phosphatidic acid, which is then primed by reacting with CTP to form CDP-diacylglycerol. This in turn reacts with inositol to re-form phosphatidylinositol. The inositol is the end product of an inositol phosphate cycle responsible for inactivating InsP_3 . The cycle begins with an inositol trisphosphatase⁴⁹, which attenuates the second messenger activity of InsP_3 by removing phosphate from the 5-position to produce inositol 1,4-bisphosphate (InsP_2) (Figs 1, 2f). An inositol bisphosphatase^{22,50} (Fig. 2g) hydrolyses InsP_2 to inositol 1-phosphate (InsP_1). In addition to inositol 1-phosphate, brain extracts also contain a significant amount of inositol 4-phosphate⁵¹, suggesting that the bisphosphatase may not be completely specific, or that there are two separate enzymes. InsP_1 is finally converted

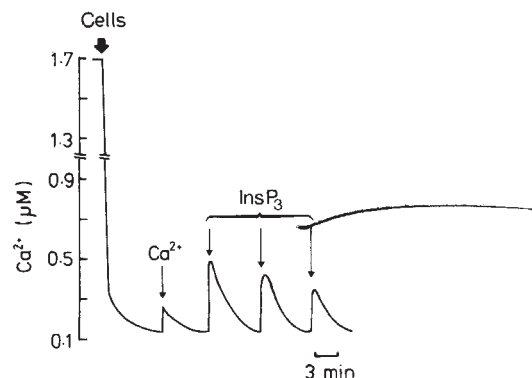


Fig. 3 Mobilization of calcium from permeabilized insulinoma cells by repeated pulses of 2.5 μM InsP_3 (ref. 66).

to free inositol by an inositol 1-phosphatase (Fig. 2h) that is markedly inhibited by lithium^{52,53}. Lithium may desensitize receptors by blocking the inositol phosphate cycle, so reducing the supply of the inositol lipids required by the receptor mechanism⁵³. In adrenal glomerulosa cells, lithium blocks the action of angiotensin II, which operates through the inositol lipids⁵⁴, but has no effect on the action of adrenal corticotrophic hormone (ACTH), which works through cyclic AMP⁵⁵.

In summary, there are four biochemical cycles concerned with the role of the inositol lipids in signal transduction (Fig. 2): the lipid and inositol phosphate cycles ultimately unite to synthesize phosphatidylinositol, which is fed into the phosphatidylinositol/PtdIns(4,5) P_2 futile cycles. The key event of the transduction mechanism is the phosphodiesterase hydrolysis of PtdIns(4,5) P_2 ; this represents a bifurcation point in the signal pathway in that the two products, diacylglycerol and InsP_3 , both function as second messengers. The role of diacylglycerol in controlling one arm of this bifurcating signal pathway has been recently reviewed⁵⁶; here we concentrate on the proposed role of InsP_3 as a second messenger to mobilize intracellular calcium^{4,23,57}.

Calcium mobilization

Evidence for the InsP_3 calcium-mobilizing hypothesis has been obtained by studying the effect of this putative second messenger on various permeabilized cells where InsP_3 could gain access to the intracellular calcium stores; this calcium-releasing property of InsP_3 was first demonstrated in a preparation of rat pancreatic acinar cells permeabilized by incubation in a low-calcium medium⁵⁷. This basic observation has been verified using different permeabilization and calcium-measuring techniques in a number of cell types including liver⁵⁸⁻⁶¹, GH₃ cells⁶², Swiss 3T3 cells^{63,64}, an insulinoma cell line^{65,66}, human neutrophils⁶⁷, macrophages¹³⁹, cardiac microsomes¹³⁹ and porcine artery smooth muscle cells⁶⁸.

For example, when permeabilized insulinoma cells are incubated with ATP in a low calcium medium (1.7 μM), they begin to sequester calcium. The result is a rapid decline in the ambient concentration towards an equilibrium value of 0.1 μM , at which point the uptake and release of calcium across the internal stores are exactly balanced⁶⁶ (Fig. 3). The dynamic nature of the system is demonstrated by the restoration of this equilibrium value following the addition of a small test pulse of calcium. Addition of InsP_3 to give a final concentration of 2.5 μM results in a rapid release of calcium followed by a slower re-uptake to the original equilibrium value⁶⁶. Similar responses are obtained following subsequent additions of InsP_3 , suggesting that the InsP_3 -sensitive calcium pool is refilled during the re-uptake phase. Identical observations have been reported for permeabilized liver cells where the re-uptake of calcium was found to coincide with the degradation of InsP_3 (ref. 59).

The intracellular calcium store appears to consist of a vesicular pool; this has been demonstrated by using the ionophores A23187 and ionomycin, which abolish the effects of InsP_3 by discharging the stores^{57,59,66,67}. The two most likely candidates

for the store are the mitochondria and the endoplasmic reticulum. These can be distinguished from each other by loading with calcium at different concentrations or by using various metabolic inhibitors. When calcium is buffered at 180 nM, permeabilized liver cells accumulate calcium in a non-mitochondrial pool which is sensitive to InsP_3 (ref. 58). Conversely, there is no release when cells are allowed to accumulate calcium in a mitochondrial pool^{58,59,66}. Inhibition of mitochondrial function by removing metabolic substrates or by adding inhibitors (antimycin, oligomycin, Ruthenium red) has no effect on the ability of InsP_3 to release calcium from the ATP-dependent vesicular pool^{57,59,62,67}. Thus, the InsP_3 -sensitive release site is probably the endoplasmic reticulum, not the mitochondria.

This conclusion is supported by cell fractionation experiments, which show that InsP_3 can release calcium from microsomes but not from isolated mitochondria^{65,69}. After further fractionation of pancreatic microsomes, InsP_3 -sensitive release is localized predominantly in those vesicles containing markers for the endoplasmic reticulum rather than the plasma membrane⁶⁹. Insulinoma microsomes do not respond to a second challenge with InsP_3 (ref. 65), which can probably be explained by vesicle heterogeneity⁶⁰. When InsP_3 is first added, calcium is chased out of the InsP_3 -sensitive vesicles only to be mopped up by those which are insensitive⁶⁰. This heterogeneity suggests that, within the intact cell, only a part of the endoplasmic reticulum, perhaps that lying close to the plasma membrane, may be sensitive to InsP_3 . Changes in intracellular calcium in response to agonists may, therefore, not be uniform, but may be localized in specific regions depending on the juxtaposition of surface receptors and intracellular calcium stores¹⁷. An analogous arrangement may occur in cardiac muscle, where calcium is thought to be taken up over the whole surface of the sarcoplasmic reticulum but to be released only from the terminal cisternae close to the T-tubules⁷⁰.

Previous studies with intact liver suggested that the hormone-dependent mobilization of calcium involves an amplification step, for more than 10 calcium ions are released for each molecule of $\text{PtdIns}(4,5)\text{P}_2$ hydrolysed by the receptor mechanism⁷¹. One amplification step probably resides at the level of the endoplasmic reticulum; for example, using data from permeabilized liver cells⁵⁹, we calculate that each InsP_3 molecule releases at least 20 calcium ions. Just how this calcium is released from the endoplasmic reticulum is unknown, but current evidence favours the view that InsP_3 acts by stimulating calcium efflux (as shown in Fig. 1) rather than by inhibiting calcium uptake. When the latter is blocked completely, either by adding vanadate or by removing ATP, there is a gradual release of calcium which is very much slower than that obtained by adding InsP_3 (ref. 67). Furthermore, the calcium mobilizing effect of InsP_3 was greatly enhanced when calcium uptake was blocked. The ability of InsP_3 to release calcium from the endoplasmic reticulum was unaffected by nifedipine (5 μM), verapamil (10 μM), cobalt (100 μM), dantrolene (30 μM) or Ruthenium red (20 μM)⁶⁶.

The specificity of the putative InsP_3 receptor on the endoplasmic reticulum (R_2 in Fig. 1) has been analysed by testing a range of inositol phosphates and closely related compounds (Table 2). It is of particular interest that InsP_2 , formed when InsP_3 is hydrolysed⁴⁹ by inositol 1,4,5-trisphosphate 5-phosphatase (Fig. 1), has no effect. The InsP_3 -binding site on that enzyme is apparently distinct from the receptor on the endoplasmic reticulum, because 2,3-diphosphoglyceric acid, which blocks the trisphosphatase⁴⁹, has no effect on the ability of InsP_3 to release calcium⁶⁴. A common feature of the four inositol phosphates which can release calcium^{61,64} (Table 2) is that they all have the pair of vicinal phosphates at the 4- and 5-positions (Fig. 1). These two phosphates appear to be essential for full activity (ability to release calcium), whereas the remaining phosphate, at the 1-position, seems to enhance the affinity of InsP_3 for its receptor. Moving this phosphate from the 1- to the 2-position causes some reduction of affinity, and removing it altogether causes a dramatic reduction in the ability of the molecule to

Table 2 Specificity of InsP_3 -sensitive calcium mobilization

		Concentrations tested (μM)
Inactive compounds		
2,3-Diphosphoglycerate ⁶⁴		2,000
Fructose 1,6-bisphosphate ⁶³		10
myo-Inositol ^{57,58,63,65,66}		2.5–100
Inositol 1-phosphate ^{57,58,65}		5–10
Inositol 2-phosphate ^{63,66,67}		7.5–10
Inositol 1,2-cyclic phosphate ^{57,58,63}		5–10
Inositol 1,4-bisphosphate ^{57,58,63–65,67}		1–50
		Concentration giving half-maximal release (μM)
Active compounds		
Inositol 1,4,5-trisphosphate	0.4	Rat pancreas acinar cells ⁵⁷
Inositol 1,4,5-trisphosphate	0.1	Rat hepatocytes ⁵⁹
Inositol 1,4,5-trisphosphate	0.2	Guinea pig hepatocytes ⁵⁸
Inositol 1,4,5-trisphosphate	0.3	Swiss 3T3 ^{63,64}
Inositol 1,4,5-trisphosphate	0.5	Rat insulinoma ⁶⁶
Inositol 1,4,5-trisphosphate	0.7	Porcine artery cells ⁶⁸
Inositol 1,4,5-trisphosphate	0.75	Human neutrophils ⁶⁷
Inositol 1,4,5-trisphosphate	0.8	Macrophages ¹³⁹
Inositol 1,4,5-trisphosphate	1.0	GH ₃ cells ⁶²
Inositol 1,4,5-trisphosphate	0.50	Rat liver microsomes ⁶⁰
Inositol 1,4,5-trisphosphate	3.0	Insulinoma microsomes ⁶⁵
Glycerophosphoinositol 4,5-bisphosphate	1.1	Swiss 3T3 ⁶⁴
Inositol 2,4,5-trisphosphate	1.1	Swiss 3T3 ⁶⁴
Inositol 2,4,5-trisphosphate	1.3	Guinea pig hepatocytes ⁶¹
Inositol 4,5-bisphosphate	20	Swiss 3T3 ⁶⁴
Inositol 4,5-bisphosphate	~50	Guinea pig hepatocytes ⁶¹

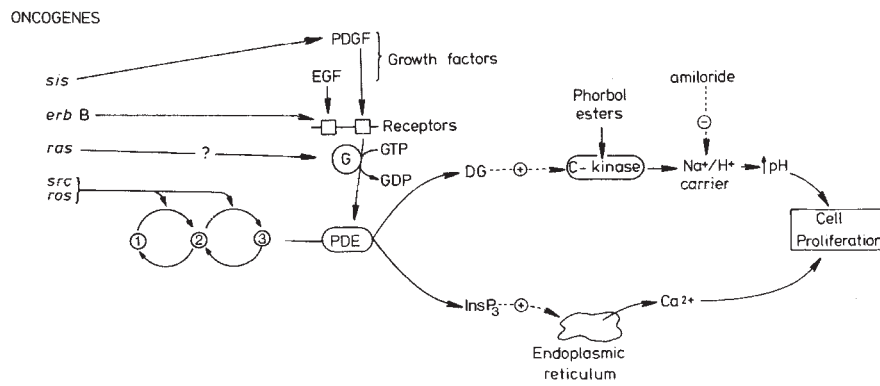
mobilize calcium (Table 2). The most effective molecule tested so far is one of the naturally occurring isomers¹⁹, with the three phosphates in a D-1,4,5-configuration.

The rate at which InsP_3 is formed after cell stimulation^{23,24,28,29} is fast enough for it to be the messenger causing calcium mobilization, but the concentration which it can reach in stimulated tissues is at present unknown. Rough estimates based on the $\text{PtdIns}(4,5)\text{P}_2$ content of tissues⁵⁷ and direct measurements of labelled InsP_3 (refs 28, 58, 72) give values in the micromolar range, which are within the levels active on permeabilized cells (Table 2). These measurements, however, all rely on isotopic techniques where the problems of pool equilibration and the existence of two isomers make accurate estimates impossible.

Phototransduction

Photoreceptors function as transducers by converting the energy of light into the electrical signals that convey visual information to the brain. A gap in our knowledge of phototransduction is the link between photon capture by rhodopsin, and the change in ion permeability that initiates electrical activity. The need for a second messenger in phototransduction is highlighted by the fact that a single photon can open approximately 1,000 sodium channels in *Limulus* photoreceptors⁷³. That InsP_3 may be such a link in phototransduction^{42,73} is borne out by evidence that light causes a decrease^{42,43} in the amount of ³H-inositol in the precursor $\text{PtdIns}(4,5)\text{P}_2$, with a corresponding increase in the amount of labelled InsP_3 (ref. 42). Octopus photoreceptors possess a protein whose GTP binding and hydrolysis functions are sensitive to light⁴¹. As described earlier (Fig. 1), this GTP-binding protein may be responsible for coupling activated receptors (in this case bleached rhodopsin) to the $\text{PtdIns}(4,5)\text{P}_2$ phosphodiesterase which generates InsP_3 .

Fig. 4 The proposed role of diacylglycerol and InsP_3 as intracellular second messengers in regulating cell proliferation. Oncogenes appear to code for various aspects of the signal pathway such as growth factors (*sis*) and the kinases (*src* and *ros*) responsible for forming the substrate ($\text{PtdIns}(4,5)\text{P}_2$) used by the receptor mechanism. The *erb-B* gene codes for a truncated EGF receptor reported to stimulate phosphatidylinositol turnover by some workers^{87,137} but not by others¹³⁸. The *ras* gene product may function to couple receptors to underlying transduction mechanisms such as the $\text{PtdIns}(4,5)\text{P}_2$ phosphodiesterase (PDE). Diacylglycerol apparently acts by raising pH, whereas InsP_3 mobilizes intracellular calcium. 1, Phosphatidylinositol; 2, $\text{PtdIns}(4)\text{P}$; 3, $\text{PtdIns}(4,5)\text{P}_2$.



When injected into the ventral photoreceptors of *Limulus*, InsP_3 induces electrical events which are identical with those produced normally by light^{42,73}. InsP_3 stimulates an inward current, carried mainly by sodium ions, which has a reversal potential identical with that induced by light. As well as functioning as a second messenger in phototransduction, InsP_3 may also be responsible for the process of adaptation^{42,73}. Injection of InsP_3 into photoreceptors reduces their responsiveness to subsequent flashes of light. Conversely, a flash of light causes a reduction in the responses induced by repetitive injections of InsP_3 (ref. 42). As an increase in intracellular calcium has been implicated in adaptation⁷⁴, it is reasonable to suppose that some of the actions of InsP_3 in photoreceptors may be mediated by mobilizing calcium from internal stores. It remains to be seen whether or not InsP_3 is involved in vertebrate photoreceptors.

Cell growth and oncogenes

The discovery of InsP_3 as an intracellular second messenger may contribute to our understanding of those aspects of cell growth and development that depend on the mobilization of calcium from intracellular stores. It has been known for some time that growth factors cause changes in inositol lipid metabolism⁷⁵, but the link between these events at the plasma membrane and the onset of cell proliferation has not been understood. There is now much support for the proposal that both diacylglycerol and InsP_3 may function as second messengers controlling the two major ionic events implicated in cell proliferation (Fig. 4). Diacylglycerol acts within the plane of the membrane to stimulate protein kinase C⁵⁶, which is probably the site of action of the tumour-promoting phorbol esters⁷⁶. As the phorbol esters can activate the Na^+/H^+ exchange carrier^{77,78}, it has been proposed that diacylglycerol may act in a similar way to increase cytoplasmic pH (refs 4, 63). Such an increase in pH is a characteristic feature of the action of growth factors and is thought to have an important role in stimulating cell growth⁷⁹⁻⁸¹.

Calcium has long been implicated in cell proliferation^{82,83}. The increase in intracellular calcium which occurs when cells are fertilized⁹ or stimulated with growth factors^{10,84} may depend on the hydrolysis of $\text{PtdIns}(4,5)\text{P}_2$ to form InsP_3 , which then releases calcium from intracellular stores. The large increase in polyphosphoinositide metabolism in sea urchin eggs following fertilization has been implicated in calcium mobilization⁸⁵. Direct evidence for the involvement of InsP_3 is provided by the observations that injection of InsP_3 (but not InsP_2) triggers some of the initial events of the fertilization process (elevation of the fertilization membrane) in sea urchin⁸⁶ and starfish oocytes (M.J.B. and L. Meijer, unpublished observations). The action of growth factors (platelet-derived growth factor (PDGF), bombesin and vasopressin) on Swiss 3T3 cells results in an increased formation of diacylglycerol⁸⁷ and InsP_3 (refs 63, 88). The latter is likely to function as a second messenger to mediate the calcium-mobilizing action of growth factors⁸⁹, as it can release calcium from permeabilized Swiss 3T3 cells^{63,64}. Thus, the hydrolysis of $\text{PtdIns}(4,5)\text{P}_2$ induced by growth factors results in the activation of two signal pathways which may regulate

two key ionic events responsible for cell proliferation (Fig. 4). The increase in pH may be controlled through diacylglycerol, and the level of calcium by InsP_3 .

The enzymatic reactions associated with this bifurcating signal pathway based on the inositol lipids may represent the sites of action of certain oncogenes (Fig. 4)⁹⁰⁻⁹². The *sis* oncogene is responsible for producing PDGF^{93,94}, known to stimulate inositol lipid metabolism^{63,87}. The *erb-B* gene codes for a truncated EGF receptor⁹⁵. (EGF stimulates an immediate increase in cytoplasmic free calcium in fibroblasts^{10,84}.) Proteins encoded by the *src* and *ros* oncogenes may function as inositol lipid kinases^{90,91} to convert phosphatidylinositol into the $\text{PtdIns}(4,5)\text{P}_2$ used by the receptor mechanism (Fig. 4). It is reasonable to speculate that other oncogenes, whose functions are still unknown, may code for enzymes concerned with the remaining steps of this bifurcating signal pathway⁹². A particularly intriguing possibility is that the activated *ras* gene may transform cells by causing an uncontrolled stimulation of $\text{PtdIns}(4,5)\text{P}_2$ phosphodiesterase (Fig. 4). The normal *ras* gene can both bind and hydrolyse GTP, but on activation by a point mutation at codon 12, the resulting oncogenic protein can still bind GTP⁴⁵, although its ability to hydrolyse this nucleotide is severely impaired^{46,47}. If *ras* does play a part in this receptor mechanism, the loss of GTPase activity means that the oncogenic protein would continue to activate the formation of InsP_3 and diacylglycerol in an uncontrolled way independently of growth factors (Fig. 4).

Most of the oncogenes discovered so far have been identified through viral transformation or transfection assays and are likely to be dominant (drive the signal pathways forward by activating positive steps such as the formation of second messengers). It is conceivable, however, that there may be a set of recessive oncogenes which code for enzymes responsible for inhibitory aspects of the signal pathways such as second messenger removal. The potential targets for both dominant and recessive oncogenes are described elsewhere⁹².

Early development

Ionic currents flowing through oocytes may be responsible for segregating cytoplasmic components to establish an anterior-posterior axis during early development^{96,97}. A calcium-dependent chloride current entering the animal pole has been detected in *Xenopus* oocytes using an extracellular vibrating electrode⁹⁷. Muscarinic receptors, which are concentrated around the animal pole, respond to acetylcholine by inducing a membrane depolarization resulting from the opening of chloride channels^{98,99} themselves regulated by calcium, which is apparently released from intracellular stores by InsP_3 (ref. 100). Injection of eggs with InsP_3 results in an electrophysiological response similar to that normally induced by acetylcholine¹⁰⁰. Even in the absence of acetylcholine, eggs often display spontaneous fluctuations in chloride conductance⁹⁸ which probably account for the chloride current that enters the animal pole. These spontaneous fluctuations may result from the transient accumulation in the animal pole of the same 'channel gating substance'

normally released by acetylcholine⁹⁸ and which could well be InsP_3 . Therefore, a localized accumulation of calcium released by InsP_3 at the animal pole may be responsible for the large influx of chloride at this point. InsP_3 could thus have an important role in early development to establish an egg axis by segregating cytoplasmic components through localized calcium gradients and electric fields.

Calcium entry

In the initial seconds or minutes of cell activation by a number of agonists, the best demonstrated function of InsP_3 as a second messenger is the mobilization of calcium from intracellular stores, particularly the endoplasmic reticulum. If, however, a tissue is to maintain an elevated calcium level, some modification of plasma membrane calcium transport is necessary because of the limited size of the intracellular calcium stores. Whether or not InsP_3 is directly involved is not known, but elevated intracellular levels of calcium could be maintained by inhibition of the calcium pump on the plasma membrane^{101,102} by some mechanism regulated by changes in the level of $\text{PtdIns}(4,5)\text{P}_2$ (ref. 7). The $\text{PtdIns}(4,5)\text{P}_2$ used to generate InsP_3 is an integral part of the bilayer, where it may regulate a number of membrane properties in addition to controlling the calcium pump. The level of $\text{PtdIns}(4,5)\text{P}_2$ may regulate membrane fluidity¹⁰³ or the anchoring and transfer of proteins across the bilayer¹⁰⁴. So, as $\text{PtdIns}(4,5)\text{P}_2$ is consumed to generate second messengers, there may be localized perturbations in the bilayer which alter a number of membrane properties.

Inositol 1,3,4-trisphosphate

In several tissues, InsP_3 levels elevated in response to agonists are prolonged over many minutes^{28,58,105,106}, but a complicating factor in measuring InsP_3 levels is that the predominant inositol triphosphate may not always be D-inositol(1,4,5)triphosphate but, rather, D- or L-inositol(1,3,4)triphosphate¹⁹. The origin and function of this is unknown, because the evidence for its existence comes from radiolabelling studies with ³H- or ¹⁴C-labelled inositol, its abundance in mass terms compared with the 1,4,5-isomer is also unknown. The occurrence of inositol(1,3,4)triphosphate is therefore a further complication in assessing the rate and extent of InsP_3 accumulation. However, there is also a possibility that inositol(1,3,4)triphosphate has messenger functions of its own and represents an entirely new facet of inositol function¹⁹. For example, hyperstimulation of the parotid via the cholinergic parasympathetic pathway, which results in a large accumulation of inositol(1,3,4)triphosphate¹⁹, also stimulates DNA synthesis¹⁰⁷. It is also significant that treatment of cells with 5–20 mM lithium, which enhances the formation of inositol triphosphate^{27,28,72} (much of which is likely to be the 1,3,4-isomer), is a potent growth stimulus in lymphocytes¹⁰⁸ and mammary¹⁰⁹ and kidney epithelial cells¹¹⁰. Although the 1,4,5-isomer seems to have more immediate effects, for example, calcium mobilization, the 1,3,4-isomer may contribute to longer-term responses, such as the onset of proliferation.

Second-messenger interactions

An important aspect of the bifurcating receptor mechanism which utilizes InsP_3 and diacylglycerol as second messengers is that the two limbs of the signal pathway may sometimes interact with each other synergistically⁵⁶, although an inhibitory effect of protein kinase C on calcium transport has also been reported¹¹¹. One explanation for the synergistic interactions is that the diacylglycerol-sensitive pathway may alter the sensitivity of the calcium pathway. For example, activation of protein kinase C by a phorbol ester results in a large increase in the calcium sensitivity of secretion in blood platelets¹¹² and adrenal medulla¹¹³. Some cellular processes, such as shape change and secretion in blood platelets, can be activated through the C-kinase pathway without a change in the resting level of calcium¹¹⁴, but a more effective stimulus is provided when both pathways act in concert^{56,112,114,115}. Another important interac-

tion is the finding that stimulation of the protein kinase C pathway can greatly enhance the formation of $\text{PtdIns}(4)\text{P}$ and $\text{PtdIns}(4,5)\text{P}_2$ (refs 116, 117). Diacylglycerol could thus function as a positive feedback signal to increase the formation of the $\text{PtdIns}(4,5)\text{P}_2$ required by the receptor mechanism^{116,117}.

Interactions between the cyclic nucleotide and inositol pathways are a relatively unexplored but potentially exciting area. In platelets, for example, cyclic AMP has a marked inhibitory effect on inositol lipid breakdown⁵⁶ (generally interpreted as an inhibition of the phosphodiesterase, but equally likely to be an effect on the kinases or phosphomonoesterases which control $\text{PtdIns}(4,5)\text{P}_2$ levels (Fig. 2)). Reciprocal effects of phorbol esters (presumably acting as diacylglycerol analogues) on the sensitivity of receptors which activate adenylate cyclase have also been reported^{118,119}.

The identification of InsP_3 as a second messenger raises far more questions than it answers. It is clear, however, that the bifurcating signal transduction system of inositol metabolism is used by a wide variety of agonists as a uniquely flexible signalling system. A key role for InsP_3 is to mobilize calcium, but its overall metabolic properties place it in a central position to carry out additional second messenger functions. The extent of its influence is evident from the growing list of tissues in which InsP_3 is a second messenger.

Note added in proof: In contrast to the lack of any effect of pertussis toxin on renal calcium-mobilizing receptors³⁶, this toxin inhibits GTP-dependent histamine release in mast cells¹⁴⁰.

- Hokin, M. R. & Hokin, L. E. *J. biol. Chem.* **203**, 967–977 (1953).
- Michell, R. H. *Biochim. biophys. Acta* **415**, 81–147 (1975).
- Downes, C. P. & Michell, R. H. *Cell Calcium* **3**, 467–502 (1982).
- Berridge, M. J. *Biochem. J.* **220**, 345–360 (1984).
- Durell, J., Garland, J. T. & Friedel, R. O. *Science* **165**, 862–866 (1969).
- Cockcroft, S. *Trends in pharmac. Sci.* **2**, 340–342 (1982).
- Hawthorne, J. N. *Biosci. Rep.* **3**, 887–904 (1983).
- Michell, R. H. & Kirk, C. J. *Trends in pharmac. Sci.* **2**, 86–90 (1981).
- Jaffe, L. F. *Dev. Biol.* **99**, 265–276 (1983).
- Moolenaar, W. H., Tertoolen, L. G. J. & de Laat, S. W. *J. biol. Chem.* **259**, 8066–8069 (1984).
- Schulz, I. & Stolze, H. H. A. *Rev. Physiol.* **42**, 127–156 (1980).
- Putney, J. W., Poggiali, J. & Weiss, S. J. *Phil. Trans. R. Soc. B296*, 37–45 (1981).
- Schlegel, W. & Wollheim, C. B. *J. Cell Biol.* **99**, 83–87 (1984).
- Hallam, T. J., Sanchez, A. & Rink, T. J. *Biochem. J.* **218**, 819–827 (1984).
- Wollheim, C. B. & Sharp, G. W. G. *Physiol. Rev.* **61**, 914–973 (1981).
- Blackmore, P. F., Hughes, B. P., Schuman, E. A. & Exton, J. H. *J. biol. Chem.* **257**, 190–197 (1984).
- Brown, R. D., Berger, K. D. & Taylor, P. J. *J. biol. Chem.* **259**, 7554–7562 (1984).
- Irvine, R. F. *Cell Calcium* **3**, 295–309 (1982).
- Irvine, R. F., Letcher, A. J., Lander, D. J. & Downes, C. P. *Biochem. J.* **223**, 237–243 (1984).
- Abdel-Latif, A. A., Akhtar, R. A. & Hawthorne, J. N. *Biochem. J.* **162**, 61–73 (1977).
- Akhtar, R. A. & Abdel-Latif, A. A. *Biochem. J.* **192**, 783–791 (1980).
- Michell, R. H., Kirk, C. J., Jones, L. M., Downes, C. P. & Creba, J. A. *Phil. Trans. R. Soc. B296*, 123–137 (1981).
- Berridge, M. J. *Biochem. J.* **212**, 849–858 (1983).
- Berridge, M. J., Buchan, P. B. & Heslop, J. P. *Molec. cell. Endocr.* **36**, 37–42 (1984).
- Rebecchi, M. J. & Gershengorn, M. C. *Biochem. J.* **216**, 299–308 (1983).
- Martin, T. F. J. *J. biol. Chem.* **258**, 14816–14822 (1983).
- Drummond, A. H., Bushfield, M. & Macphee, C. H. *Molec. Pharmac.* **25**, 201–208 (1984).
- Thomas, A. P., Alexander, J. & Williamson, J. R. *J. biol. Chem.* **259**, 5574–5584 (1984).
- Drummond, A. H., Knox, R. J. & Macphee, C. H. *Biochem. Soc. Trans.* (in the press).
- Seyfred, M. A. & Wells, W. W. J. *J. biol. Chem.* **259**, 7666–7672 (1984).
- Irvine, R. F., Letcher, A. J. & Dawson, R. M. C. *Biochem. J.* **218**, 177–185 (1984).
- Gomperts, B. D. *Nature* **306**, 64–66 (1983).
- Gilman, A. G. *Cell* **36**, 577–579 (1984).
- Goodhardt, M., Ferry, N., Geynet, P. & Hanoune, J. *J. biol. Chem.* **257**, 11577–11583 (1982).
- Hulme, E. C., Berrie, C. P., Birdsall, N. J. M. & Burgen, A. S. V. in *Drug Receptors and their Effectors* (ed. Birdsall, N. J. M.) p 23–34 (Macmillan, London, 1981).
- Boyer, J. L., Garcia, A., Posadas, C. & Garcia-Sainz, J. A. *J. biol. Chem.* **259**, 8076–8079 (1984).
- Haslam, R. J. & Davison, M. M. L. *FEBS Lett.* **174**, 90–95 (1984).
- Haslam, R. J. & Davidson, M. M. L. *J. Receptor Res.* (in the press).
- Fein, A. & Corson, D. W. *Science* **212**, 555–557 (1982).
- Bolsover, S. R. & Brown, J. E. *J. Physiol., Lond.* **332**, 325–342 (1982).
- Vandenberg, C. A. & Montal, M. *Biochemistry* **23**, 2339–2347 (1984).
- Brown, J. E. *et al. Nature* **311**, 160–162 (1984).
- Vandenberg, C. A. & Montal, M. *Biochemistry* **23**, 2347–2352 (1984).
- Bishop, J. M. A. *Rev. Biochem. Sci.* **52**, 301–354 (1983).
- Finkel, T., Der, C. J. & Cooper, G. M. *Cell* **37**, 151–158 (1984).
- McGrath, J. P., Capon, D. J., Goeddel, D. V. & Levinson, A. D. *Nature* **310**, 644–649 (1984).
- Sweet, R. W. *et al. Nature* **311**, 273–275 (1984).
- Kamata, T. & Feramisco, J. R. *Nature* **310**, 147–150 (1984).
- Downes, C. P., Mussat, M. C. & Michell, R. H. *Biochem. J.* **203**, 169–177 (1982).
- Berridge, M. J., Dawson, R. M. C., Downes, C. P., Heslop, J. P. & Irvine, R. F. *Biochem. J.* **212**, 473–482 (1983).
- Sherman, W. R., Munsell, L. Y., Gish, B. G. & Honchar, M. P. J. *Neurochem.* (in the press).
- Hallcher, L. M. & Sherman, W. R. *J. biol. Chem.* **255**, 10896–10901 (1980).
- Berridge, M. J., Downes, C. P. & Hanley, M. R. *Biochem. J.* **206**, 587–595 (1982).
- Farese, R. V., Larson, R. E. & Davis, J. S. *Endocrinology* **114**, 302–304 (1984).
- Balla, T., Enyedi, P., Hunyadi, L. & Spät, A. *FEBS Lett.* **171**, 179–182 (1984).
- Nishizuka, Y. *Nature* **308**, 693–698 (1984).
- Streb, H., Irvine, R. F., Berridge, M. J. & Schulz, I. *Nature* **306**, 67–69 (1983).
- Burgess, G. M. *et al. Nature* **309**, 63–66 (1984).

59. Joseph, S. K., Thomas, A. P., Williams, R. J., Irvine, R. F. & Williamson, J. R. *J. biol. Chem.* **259**, 3077-3081 (1984).
60. Dawson, A. P. & Irvine, R. F. *Biochem. biophys. Res. Commun.* **120**, 858-864 (1984).
61. Burgess, G. M., Irvine, R. F., Berridge, M. J., McKinney, J. S. & Putney, J. W. *Biochem. J.* (in the press).
62. Gershenhorn, M. C., Geras, E., Purrello, V. S. & Rebecchi, M. J. *J. biol. Chem.* **259**, 10675-10681 (1984).
63. Berridge, M. J., Heslop, J. P., Irvine, R. F. & Brown, K. D. *Biochem. J.* **222**, 195-201 (1984).
64. Irvine, R. F., Brown, K. D. & Berridge, M. J. *Biochem. J.* **222**, 269-272 (1984).
65. Prentki, M. *et al. Nature* **309**, 562-564 (1984).
66. Biden, T. J., Prentki, M., Irvine, R. F., Berridge, M. J. & Wollheim, C. B. *Biochem. J.* **223**, 467-475 (1984).
67. Prentki, M., Wollheim, C. B. & Lew, P. D. *J. biol. Chem.* (in the press).
68. Suematsu, E., Hirata, M., Hashimoto, T. & Kuriyama, H. *Biochem. biophys. Res. Commun.* **120**, 481-485 (1984).
69. Streb, H., Bayerdorffer, E., Haase, W., Irvine, R. F. & Schulz, I. *J. Membrane Biol.* (in the press).
70. Jorgensen, A. O., Kalnins, V. & MacLennan, D. H. *J. Cell Biol.* **80**, 372-384 (1979).
71. Creba, J. A. *et al. Biochem. J.* **212**, 733-747 (1983).
72. Rubin, R. P., Godfrey, P. P., Chapman, D. A. & Putney, J. W. *Biochem. J.* **219**, 655-659 (1984).
73. Fein, A., Payne, R., Corson, D. W., Berridge, M. J. & Irvine, R. F. *Nature* **311**, 157-160 (1984).
74. Lisman, J. E. & Brown, J. E. *J. gen. Physiol.* **59**, 701-719 (1972).
75. Mitchell, R. H. *Cell Calcium* **3**, 429-440 (1982).
76. Castagna, M. *et al. J. biol. Chem.* **257**, 7847-7851 (1982).
77. Burns, C. P. & Rozengurt, E. *Biochem. biophys. Res. Commun.* **116**, 931-938 (1983).
78. Rosoff, P. M., Stein, L. F. & Cantley, L. C. *J. biol. Chem.* **259**, 7056-7060 (1984).
79. Moolenaar, W. H., Tsein, R. Y., van der Saag, P. T. & de Laat, S. W. *Nature* **304**, 645-648 (1983).
80. Cassel, D., Rothenberg, P., Zhuang, Y.-X., Deuel, T. F. & Glaser, L. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6224-6228 (1983).
81. Lopez-Rivas, A., Stroobant, P., Waterfield, M. D. & Rozengurt, E. *EMBO J.* **3**, 939-944 (1984).
82. Boynton, A. L., Whitfield, J. F. & Isaacs, R. J. *In Vitro* **10**, 12-17 (1974).
83. Berridge, M. J. *J. cyclic Nucleotide Res.* **1**, 305-320 (1975).
84. Morris, J. D. H., Metcalfe, J. C., Smith, G. A., Hesketh, T. R. & Taylor, M. V. *FEBS Lett.* **169**, 189-193 (1984).
85. Turner, P. R., Sheetz, M. P. & Jaffe, L. A. *Nature* **310**, 414-419 (1984).
86. Whitaker, M. & Irvine, R. F. *Nature* (in the press).
87. Habenicht, A. J. R. *et al. J. biol. Chem.* **256**, 12329-12335 (1981).
88. Brown, K. D., Blay, J., Irvine, R. F., Heslop, J. P. & Berridge, M. J. *Biochem. biophys. Res. Commun.* **123**, 377-384 (1984).
89. Lopez-Rivas, A. & Rozengurt, E. *Biochem. biophys. Res. Commun.* **114**, 240-247 (1983).
90. Sugimoto, Y., Whitman, M., Cantley, L. C. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2117-2121 (1984).
91. Macara, I. G., Marinetti, G. V. & Balduzzi, P. C. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2728-2732 (1984).
92. Berridge, M. J. *Biotechnology* **2**, 541-546 (1984).
93. Doolittle, R. F. *et al. Science* **221**, 275-277 (1983).
94. Waterfield, M. D. *et al. Nature* **304**, 35-39 (1983).
95. Downward, J. *et al. Nature* **307**, 521-527 (1984).
96. Jaffe, L. F. in *Membrane Transduction Mechanisms* (eds Cone, R. A. & Dowling, J. E.) 199-231 (Raven, New York, 1979).
97. Robinson, K. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 837-841 (1979).
98. Kusano, K., Miledi, R. & Stinnakre, J. *J. Physiol., Lond.* **328**, 147-170 (1982).
99. Dascal, N., Landau, E. N. & Lass, Y. *J. Physiol., Lond.* **352**, 551-574 (1984).
100. Oron, Y., Dascal, N., Nadler, E. & Lupu, M. *Nature* (in the press).
101. Prpic, V., Green, K. C., Blackmore, P. F. & Exton, J. H. *J. biol. Chem.* **259**, 1382-1385 (1984).
102. Lin, S.-H., Wallace, M. A. & Fain, J. N. *Endocrinology* **113**, 2268-2275 (1983).
103. Sheetz, M. P., Febrriello, P. & Koppel, D. E. *Nature* **296**, 91-93 (1982).
104. Donovan, J. J., Simon, M. I. & Montal, M. *Nature* **298**, 669-672 (1982).
105. Downes, C. P. & Wusteman, M. M. *Biochem. J.* **216**, 633-640 (1983).
106. Aub, D. L. & Putney, J. W. *Life Sci.* **34**, 1347-1355 (1984).
107. Schneyer, C. A. in *Secretory Mechanisms of Exocrine Glands* (eds Thorn, N. A. & Petersen, O. H.) 42-55 (Munksgaard, Copenhagen, 1974).
108. Hart, D. A. *Expl Cell Res.* **119**, 47-53 (1979).
109. Tomooka, Y., Imagawa, W., Nandi, S. & Bern, H. A. *J. cell. Physiol.* **117**, 290-296 (1983).
110. Ryback, S. M. & Stockdale, F. E. *Expl Cell Res.* **136**, 263-270 (1981).
111. Lagast, H., Pozzan, T., Waldvogel, F. A. & Lew, P. D. *J. clin. Invest.* **73**, 878-888 (1984).
112. Knight, D. E. & Scrutton, M. C. *Nature* **309**, 66-68 (1984).
113. Knight, D. E. & Baker, P. F. *FEBS Lett.* **160**, 98-100 (1983).
114. Rink, T. J., Sanchez, A. & Hallam, T. J. *Nature* **305**, 317-319 (1983).
115. Haslam, R. J. & Davidson, M. M. *L. Biochem. J.* **222**, 351-361 (1984).
116. De Chaffoy De Courcelles, D., Roevens, P. & Van Belle, H. *FEBS Lett.* **173**, 389-393 (1984).
117. Taylor, M. V., Metcalfe, J. C., Hesketh, T. R., Smith, G. A. & Moore, J. P. *Nature* (in the press).
118. Sibley, D. R., Nambi, P., Peters, J. R. & Lefkowitz, R. J. *Biochem. biophys. Res. Commun.* **121**, 973-979 (1984).
119. Kelleher, D. J., Pressin, J. E., Ruoho, A. E. & Johnson, G. L. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4316-4320 (1984).
120. Fisher, S. K., Hootman, S. R., Heacock, A. M., Ernst, S. A. & Agranoff, B. W. *FEBS Lett.* **155**, 43-46 (1983).
121. Shukla, S. D., Buxton, D. B., Olson, M. S. & Hanahan, D. J. *J. biol. Chem.* **258**, 10212-10214 (1983).
122. Litosch, I., Lee, H. S. & Fain, J. N. *Am. J. Physiol.* **246**, C141-C147 (1984).
123. Leung, N. L., Vickers, J. D., Kinlough-Rathbone, R. L., Reimers, H.-J. & Mustard, J. F. *Biochem. biophys. Res. Commun.* **113**, 483-490 (1983).
124. Agranoff, B. W., Murthy, P. & Seguin, E. B. *J. biol. Chem.* **258**, 2076-2078 (1983).
125. Billah, M. M. & Lapetina, E. G. *J. biol. Chem.* **257**, 12705-12708 (1982).
126. Shukla, S. D. & Hanahan, D. J. *Archs Biochem. Biophys.* **227**, 626-629 (1983).
127. Bone, E. A., Fretten, P., Palmer, S., Kirk, C. J. & Michell, R. H. *Biochem. J.* **221**, 803-811 (1984).
128. Laychock, S. G. *Biochem. J.* **216**, 101-106 (1983).
129. Volpi, M., Yassin, R., Naccache, P. H. & Sha'afi, R. I. *Biochem. biophys. Res. Commun.* **112**, 957-964 (1983).
130. Yano, K., Nakashima, S. & Nozawa, Y. *FEBS Lett.* **161**, 296-300 (1983).
131. Dougherty, R. W., Godfrey, P. P., Hoyle, P. C., Putney, J. W. & Freer, R. J. *Biochem. J.* **222**, 307-314 (1984).
132. Hirayama, T. & Kato, I. *FEBS Lett.* **157**, 46-50 (1983).
133. Hasegawa-Sasaki, H. & Sasaki, T. *Biochim. biophys. Acta* **754**, 305-314 (1983).
134. Beaven, M. A., Moore, J. P., Smith, G. A., Hesketh, T. R. & Metcalfe, J. C. *J. biol. Chem.* **259**, 7137-7142 (1984).
135. Yano, K., Higashida, H., Inoue, R. & Nozawa, Y. *J. biol. Chem.* **259**, 10201-10207 (1984).
136. Irvine, R. F. *Biochem. J.* **204**, 3-16 (1982).
137. Sawyer, S. T. & Cohen, S. *Biochemistry* **20**, 6280-6286 (1981).
138. Macphie, C. H., Drummond, A. H., Otto, A. M. & Jimenez de Asua, L. *J. cell. Physiol.* **119**, 35-40 (1984).
139. Hirata, M., Suematsu, E., Hashimoto, T., Hamachi, T. & Koga, T. *Biochem. J.* **223**, 229-236 (1984).
140. Nakamura, T. & Ui, M. *FEBS Lett.* **173**, 414-418 (1984).

ARTICLES

Water column dissolution of aragonite in the Pacific Ocean

Robert H. Byrne*, James G. Acker*, Peter R. Betzer*, Richard A. Feely† & Marion H. Cates‡

* Department of Marine Science, University of South Florida, St Petersburg, Florida 33701, USA

† Pacific Marine Environmental Laboratory, NOAA 7600 Sand Point Way NE, Seattle, Washington 98115, USA

‡ Honeywell Inc., Mil-Av Division 13350 US Highway 19S, St Petersburg, Florida 33546, USA

Pelagic pteropods are an important component of the oceanic CO₂ system. Their shells appear to be a principal source of excess alkalinity in the upper water column of the North Pacific Ocean. Dissolution of aragonite must therefore be taken account of in generating models of the oceanic CO₂ system and is also an important cause of sample loss from long-term sediment traps deployed in undersaturated seawater.

MARINE aragonitic detritus is an important component of the oceanic CO₂ system and may constitute worldwide about one half the calcium carbonate (CaCO₃) detritus settling to the ocean floor¹; in some areas of the Pacific Ocean it appears to constitute substantially more than half the total CaCO₃ flux in the upper ocean². Aragonite dissolution provides a means of

explaining alkalinity anomalies in the upper ocean^{2,3} and, in addition, a means of balancing the oceanic CaCO₃ mass budget^{1,4}. Aragonite may be especially important in the short-term buffering of fossil fuel CO₂ injected into the ocean⁵.

The importance of aragonite in the oceans is largely attributable to its rapid dissolution kinetics compared with calcite.

Table 1 Aragonite dissolution rates

$\Omega_{\text{calcite}}^*$	Keir†	Morse <i>et al.</i> †	Eq. (8)†	Eq. (7)‡
1.50	0.09	0.36	0.30	0.46
1.30	0.9	1.9	1.7	1.6
1.15	3.0	4.2	4.2	3.5
1.00	7.5	8.1	8.2	6.6
0.85	16	14	15	12
0.75	25	18	20	16
0.65	38	25 (34)§	27	22
0.55	56	33 (65)§	36	30

Rates (% d⁻¹) predicted by Morse *et al.*⁶, Keir⁷ and this work as a function of saturation state at 5 °C, 34.03‰ salinity and pressures of 60–760 atm. In this analysis we have used the properties of a seawater sample collected at 26° N, 165° E in the North Pacific Ocean.

* Ω_{calcite} calculated using the equations and procedures in Millero¹¹.

† $\rho = 1.78$.

‡ $\rho = 2.05$.

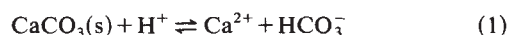
§ The higher dissolution rates shown in parentheses calculated using equation (11).

Consequently, these kinetics have received considerable attention recently. Substantial differences exist between results from careful laboratory investigations^{6,7} and *in situ* analyses^{8,9}: *in situ* aragonite dissolution rate determinations are more than an order of magnitude smaller than predictions based on laboratory analyses. These differences are a major concern in assessing the behaviour of aragonite in the oceans. We have performed a large number of dissolution rate determinations at sea, using freshly collected aragonitic particulates in an attempt to resolve the discrepancy. Our methods are only a small step removed from *in situ* analyses and incorporate advantageous features of laboratory methodology.

Analytical techniques

Our analyses were conducted aboard the RV *Discoverer*, along 165° E in the North Pacific Ocean (station locations and hydrographic data provided in Feely *et al.*¹⁰). Our pteropod samples were collected using free-drifting sediment traps at depths between 100 and 2,200 m. Use of large area (0.66 m²) traps allowed short (~24 h) collection periods and collection of intact pteropod shells with surface chemistries presumably quite similar to those of pteropods settling through the water column. The seawater used in our dissolution analyses was collected at the sediment trap stations along our transect². These samples were generally collected near 2,000 m and were characterized with respect to total CO₂, carbonate alkalinity and pH, as described previously¹⁰.

Our experiments were conducted by examining the dissolution of intact pteropod shells devoid of internal organic matter. In general, individual pteropods were used in each of our 80 experiments. Pteropods were housed in glass-plastic sample vials inside thermostated high-pressure vessels and were gently oscillated at 7 cycles min⁻¹ for the duration of each experiment. Each dissolution experiment was conducted at 5 °C and at pressures near 375 atm; extent of dissolution was monitored by measuring the seawater pH in the sample vials before and after each experiment. The pH changes induced by CaCO₃(s) dissolution



were generally quite small (0.03 ≤ pH ≤ 0.20). In order to maximize the signal-to-noise ratio, our seawater sample volumes were small (~9 ml) and pH was measured using electrodes (Ross-type reference—Orion Research) substantially free of drift in the course of thermal cycling. The total CaCO₃ dissolved in each experiment was determined using the initial alkalinity, pH and total CO₂ characteristics of our seawater samples¹⁰ and by calculating the alkalinity change (twice the change in total CO₂)

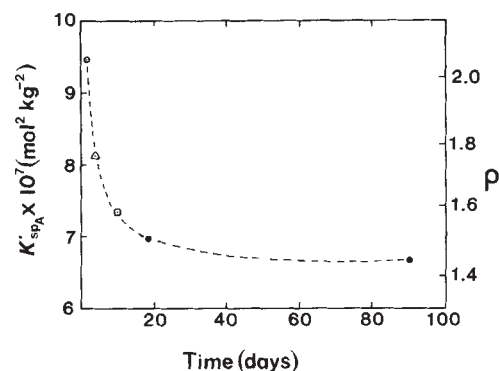


Fig. 1 The solubility behaviour of aragonite obtained in various investigations as a function of equilibration time¹⁸. The parameter $\rho = K'_{\text{spA}}/K'_{\text{spC}}$ provides a direct comparison of K'_{spC} , the solubility product of calcite obtained by Ingle *et al.*²⁰ (35‰ salinity, 25 °C) and K'_{spA} , values (35‰ salinity, 25 °C) obtained by: ○, Plath¹⁶; △, Berner¹⁵; □, MacIntyre¹⁷; ●, Morse *et al.*¹⁸. K'_{spC} values obtained in seawater at 25 °C and 35‰ salinity^{14–18,20} agree to within 5%. Consequently ρ is not significantly altered by alternative choices of K'_{spC} . Whereas the minimum value of ρ is reasonably well established, upper-bound values of ρ , appropriate to natural aragonitic particulates undergoing dissolution, are not well defined.

necessary to produce the observed change in pH. The equations and formation constants used in our calculations follow Millero¹¹ and Feely *et al.*¹⁰. Subsequent measurements of sample (pteropod) masses on laboratory microbalances permitted the result of each experiment to be expressed in terms of a % weight loss.

A potential liability encountered in the use of fresh pteropod shells is occasional evidence of bacterial respiration during dissolution experiments. We found an effective remedy for this possibility in the addition of Hg(II) to the seawater samples. Our seawater media were 2×10^{-5} M in Hg(II). Hg(II) has a very substantial affinity for halide ions, so Hg²⁺ is not readily available for interaction with carbonate surfaces. HgCl₄²⁻ substantially dominates the speciation scheme of Hg(II) in seawater¹² and, in our media, free Hg²⁺ concentrations are $\sim 10^{-19}$ M.

Detailed laboratory experiments^{6,7,13} have shown that carbonate dissolution rates in seawater generally can be well described by the following expression

$$R = k(1 - \Omega)^n \quad (2)$$

where R is the dissolution rate (% per day), k has the dimensions % per day, n is the reaction order and Ω is the degree of saturation calculated as

$$\Omega = \frac{m_{\text{Ca}} m_{\text{CO}_3}}{K'_{\text{spA}}} \quad (3)$$

where m_{Ca} and m_{CO_3} are total Ca²⁺ and CO₃²⁻ concentrations in mol kg⁻¹ seawater, and K'_{spA} is the temperature, pressure and salinity-dependent apparent solubility product of aragonite in seawater. The cumulative % dissolution in each of our experiments was modelled using the following integral form of equation (2):

$$D = \int_0^t R \, dt = k \int_0^t \left[1 - \frac{m_{\text{Ca}} m_{\text{CO}_3}}{K'_{\text{spA}}} \right]^n \, dt \quad (4)$$

Dissolution times in these experiments ranged between one hour and approximately one day. The integrations were performed numerically by dividing the total dissolution time, t , in each experiment, into a large number of segments. Our analyses indicated that dissolution segments ≤ 1 min produced identical model predictions.

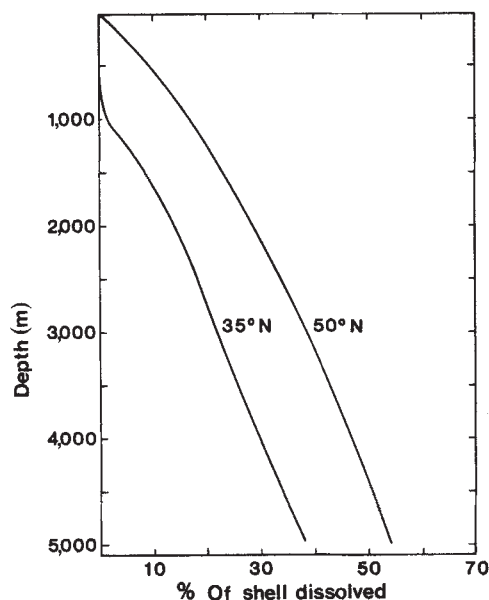


Fig. 2 The cumulative % dissolution of pteropods settling at a mid-range rate, 1.4 cm s^{-1} , at two of our station locations in the North Pacific Ocean. Significant dissolution begins at shallow depths at the northernmost station (50° N , 167° E). At 35° N , 165° E dissolution becomes appreciable below 1,000 m. As the influence of dissolution on settling rate is not considered, the cumulative dissolution shown here should be considered as a lower-bound estimate.

Equation (4) was applied here to each of our 80 dissolution experiments. In each analysis a set of parameters (k , K'_{spA} and n) was chosen and held constant. This process generated a value, $D(\text{predicted})$, for each of our experiments. $D(\text{predicted})$ was then compared with $D(\text{observed})$ and the difference was used to generate a residual, ϕ_i . The quality of a given set of parameters (k , K'_{spA} and n) was evaluated by examining the residual sum of squares generated by modelling the entire 80 experiment data set:

$$\text{residual sum of squares} = \sum_{i=1}^{80} \phi_i^2 \quad (5)$$

According to this process, the residual sum of squares was examined as a function of k , K'_{spA} and n . Our choices of K'_{spA} were constrained by the results obtained in a variety of direct aragonite solubility determinations^{14–18}. An unusual feature of aragonite solubility relative to that of calcite is the long equilibrium time necessary for attainment of true equilibrium. Morse *et al.*¹⁸ demonstrated that where the aragonite solubility products obtained in various investigations are ordered with respect to equilibration time, a monotonic decrease is observed over a period of weeks (Fig. 1). This behaviour appears to reflect slow variation in the surface chemistry of aragonite¹⁹ and indicates that K'_{spA} appropriate to naturally occurring aragonitic particulates may assume values $\geq 6.65 \times 10^{-7} \text{ (mol}^2 \text{ kg}^{-2}\text{)}$. For convenience, K'_{spA} was expressed in terms of the apparent solubility product of calcite²⁰:

$$K'_{\text{spA}} = \rho K'_{\text{spC}} \quad (6)$$

Solubility analysis¹⁴ and partial molal volume data indicate that the effects of temperature and pressure on ρ are quite small compared with the changes in ρ attributable to variations in aragonite surface chemistry. K'_{spC} was calculated according to the methods of Millero¹¹ and Feely *et al.*¹⁰; values of ρ representative of the range of $K'_{\text{spA}}/K'_{\text{spC}}$ values obtained in detailed solubility analyses were used to calculate K'_{spA} . The values of ρ used here correspond to the lowest and highest values of ρ , 1.45^{18} and 2.05^{16} respectively, at 25° C and 35% ; and to the ratio,

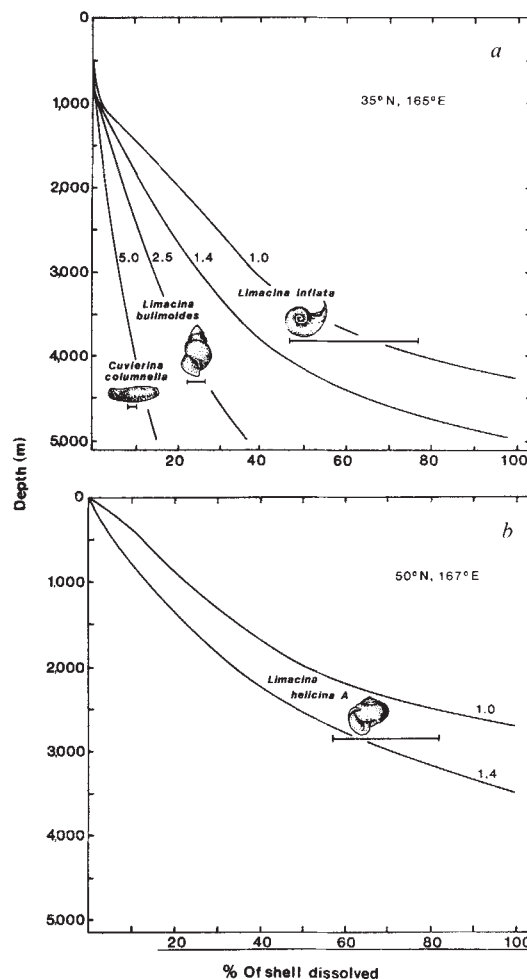


Fig. 3 a, b, The predicted cumulative % dissolution of pteropod shells in the North Pacific versus depth for representative initial settling velocities. The initial settling rates appropriate to each species are based on results obtained in our laboratory. The scales adjacent to each pteropod species represent a length of 1 mm. b, A single pteropod species was observed in the highly undersaturated waters at the northernmost station.

$\rho = 1.78$, appropriate to the use of Berner's K'_{spA} data¹⁵ in the dissolution analyses of Morse *et al.*¹⁶ and Keir⁷.

Aragonite dissolution rates

Examination of the residual sum of squares surfaces (RSS, k , n , ρ) generated by our data and the dissolution model, equation (4), produced the following results. First, the data could not be satisfactorily modelled using values of ρ as small as 1.45. Second, the best description of our results is provided by the equation

$$R = 125 \left[1 - \frac{m_{\text{Ca}} m_{\text{CO}_3}}{2.05 K'_{\text{spC}}} \right]^{4.1} \quad (7)$$

Third, the best fit model obtained with $\rho = 1.78$ produced reasonable descriptions of the data, only slightly inferior to the best fit results from equation (7)

$$R = 130 \left[1 - \frac{m_{\text{Ca}} m_{\text{CO}_3}}{1.78 K'_{\text{spC}}} \right]^{3.1} \quad (8)$$

Our dissolution rate results can be compared with the dissolution models of Morse *et al.*⁶ and Keir⁷. The dissolution rate equation of Keir⁷ is given by

$$R(\% \text{ per day}) = 318(1 - \Omega)^{4.2} \quad (9)$$

Morse *et al.*⁶ described their dissolution rate results using two

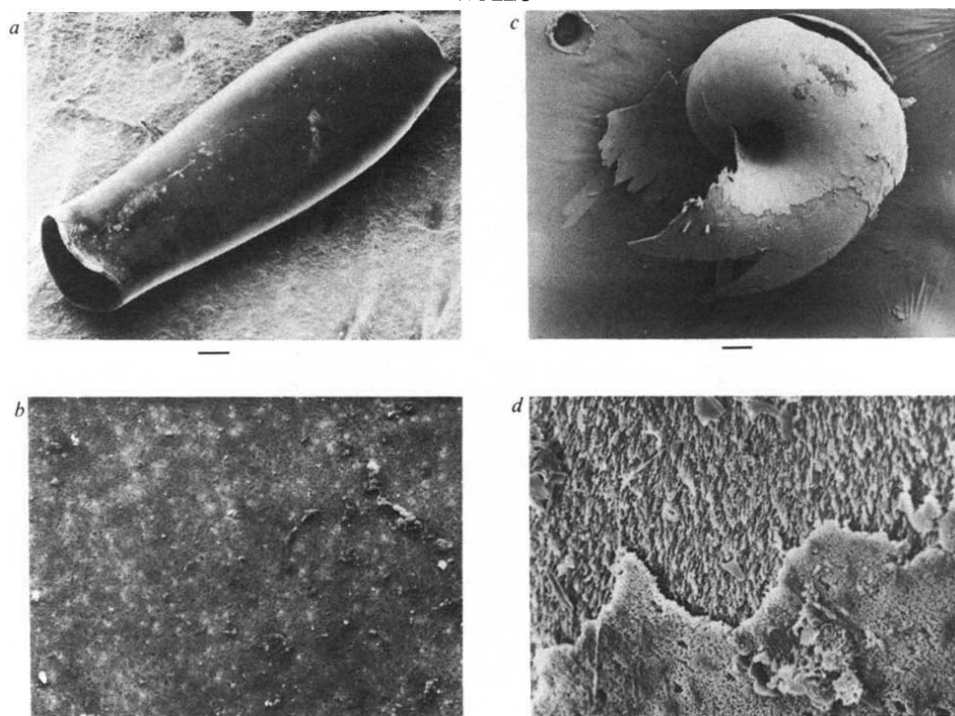


Fig. 4 The typical surface transformations accompanying pteropod dissolution. **a**, A smooth surface (*C. columnella*, scale bar 500 μm) typical of all large, rapidly settling species found in our sediment traps. **b**, Pteropods not exposed to undersaturated seawater exhibit smooth surfaces, even under high magnification (scale bar 10 μm). **c**, In contrast, the very thin-walled, slowly settling species, *L. inflata* (scale bar 100 μm) exhibits considerable surface etching. **d**, Under high magnification (scale bar 10 μm) this individual, taken from 2,100 m at 35° N, 165° E, shows the sharply contrasting surface characteristics typical of partially dissolved, slowly settling species.

equations. For $\Omega_A \geq 0.44$ they obtained the result

$$R(\% \text{ per day}) = 110(1 - \Omega)^{2.93} \quad (10)$$

and for higher degrees of undersaturation ($\Omega_A \leq 0.44$) they obtained

$$R(\% \text{ per day}) = 1,318(1 - \Omega)^{7.27} \quad (11)$$

Despite apparent differences in equations (7)–(11), the dissolution rate predictions of each are generally in good agreement. The dissolution rate predictions obtained here and by Morse *et al.*⁶ and Keir⁷ are compared (Table 1). In examining their predictions for the conditions of our experiments, we assumed a value of $\rho = 1.78$. This is consistent with their use of Berner's K'_{SPA} values.

The dissolution rates shown in Table 1 indicate that the determinations of Morse *et al.* and Keir provide good first-order predictions for the dissolution of aragonite in the oceans. This is important because previous comparisons have been either semiquantitative^{6,21} or in very poor agreement^{8,9} with laboratory results. In agreement with Morse *et al.*⁶ and Morse and Berner¹³, our results indicate that the *in situ* devices of Milliman⁸ and Honjo and Erez⁹ produced saturation states significantly exceeding the surrounding seawater.

Differences in the results obtained here and by Morse *et al.* and Keir can be attributed to a variety of factors. The enhanced dissolution rates observed by Keir are probably the result of using pteropod fragments rather than whole pteropods. The inner surfaces of intact pteropods have a somewhat limited availability for dissolution. Consequently, the experiments of Keir were conducted using materials of a somewhat larger surface area to mass ratio than the intact pteropods used here. The results of Morse *et al.*⁶ are in remarkable agreement with our results, except at high degrees of undersaturation. This indicates that our results are commensurate with equation (10) but not equation (11). This may indicate that, in the low temperature environment of the deep ocean, the change in reaction order observed at $\Omega_A = 0.44$ by Morse *et al.*⁶ occurs at much higher degrees of undersaturation. Finally, in agreement with

the solubility analyses of Morse *et al.*¹⁸, our results indicate that whereas lower bound K'_{SPA} values are reasonably well defined, the upper bound K'_{SPA} values appropriate to freshly collected pelagic aragonitic particulates are not well defined. $K'_{\text{SPA}}/K'_{\text{SPC}}$ ratios appropriate to freshly collected aragonite in undersaturated seawater appear to exceed significantly the ratio attained after long equilibration periods.

Water column dissolution of pteropods

The results given in equations (7)–(11) can be used to examine the cumulative dissolution of pteropods settling through the oceanic water column. Using the carbonate system characterizations (Ω_C versus depth) at two of our stations in the North Pacific¹⁰ we have calculated the cumulative % dissolution of pteropods settling at a mid-range rate²² of 1.4 cm s^{-1} . The results of this analysis at two locations along the cruise track are shown in Fig. 2.

At a mid-range settling rate, pteropods in transit to the floor of the Pacific Ocean will undergo substantial dissolution (Fig. 2). At high latitudes, dissolution rates are appreciable at shallow depths, whereas at low latitudes dissolution becomes appreciable at depths below 1 km. It is important to recognize that Fig. 2 was constructed without considering the decrease in settling velocity which accompanies mass loss. Therefore, we regard the extents of dissolution shown in Fig. 2 as lower bound estimates for the initial, 1.4 cm s^{-1} settling velocity.

Analysis of the influence of mass loss on settling velocity is complicated by considerable variations in pteropod morphology. Nevertheless, a reasonable first-order assessment of the effect of mass loss on settling velocity can be constructed by assuming the following relationship:

$$v \equiv v_0 \left(\frac{m}{m_0} \right) \quad (12)$$

where v is pteropod settling velocity, v_0 is initial pteropod settling velocity, m_0 is initial pteropod shell mass, and m is pteropod shell mass corresponding to velocity v . Equation (12) is appropriate where substantial mass losses occur without

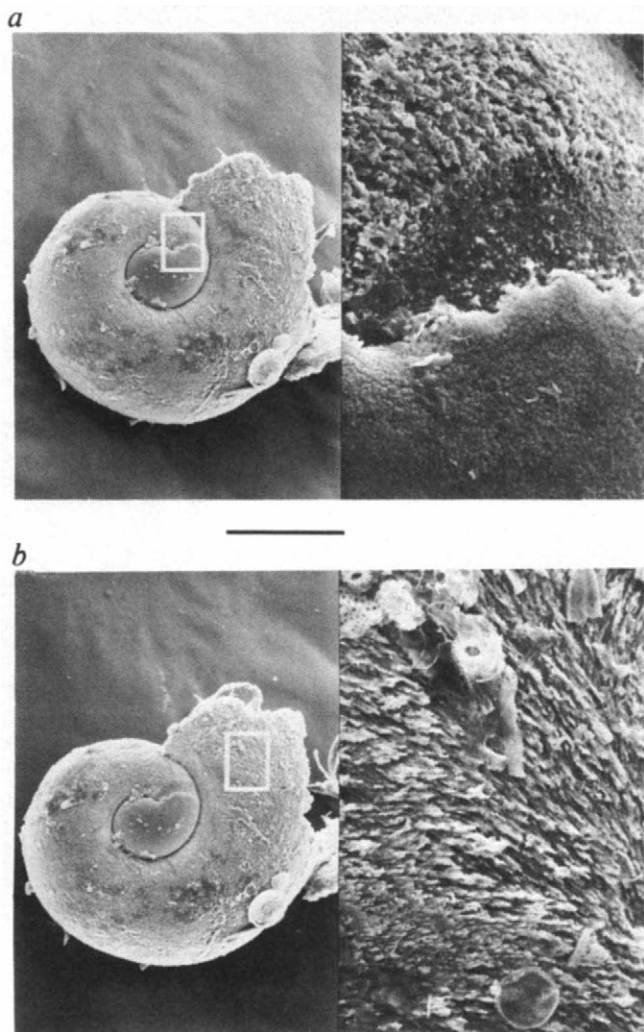


Fig. 5 Advanced stages of dissolution in smaller size class pteropods (immature *L. inflata*) found in the 2,100-m sediment traps at 35° N, 165° E. *a*, In this individual (scale bar 100 μ m) only a very small portion of the original smooth surface remains. *b*, The considerable surface roughening which occurs after extended exposure to highly undersaturated seawater.

appreciable changes in particle volume and, as such, is applicable to thin-walled shells until onset of fragmentation. Figure 3 shows the predicted dissolution behaviour of pteropods at several initial settling velocities, using equation (12) to account for reductions in settling velocity.

Discussion

Although our inability to deal with the consequences of shell fragmentation makes the predictions shown in Fig. 3 somewhat uncertain, it is clear that the influence of mass loss on settling rate can produce striking differences in the behaviour of pteropod shells in the water column. Shells settling at a mid-range rate are substantially dissolved before reaching the ocean floor. Shells settling with an initial velocity of 100 cm s^{-1} , appropriate to a major size class (juvenile *Limacina inflata*) observed in our mass flux investigations dissolve in the upper water column of the Pacific Ocean. It is important to note, however, that even in the corrosive waters found in the Pacific, large pteropods such as *Cuvierina columnella* undergo relatively minor dissolution before reaching the ocean floor. This explains the otherwise disconcerting observations that pteropods are found on the uppermost sediment of box cores²³ and as empty shells in the deep water column of the Pacific Ocean²⁴.

The results shown in Figs 2 and 3 have important implications in the interpretation of CaCO_3 flux measurements in the ocean.

Long-duration sediment trap deployments in undersaturated seawater will lead to large underestimations of aragonite fluxes. The aragonite fluxes observed by Betzer *et al.*², using short-duration sediment trap deployments, are quite substantial compared with previous results obtained through long-term collections. The flux measurements of Betzer *et al.* are lower bound estimates: large mobile pteropods which may have been able to swim into our sediment traps were not included in their estimates; because some accounting of large pteropod contributions to mass flux is certainly necessary, the significance of aragonite in the Pacific Ocean could exceed considerably the lower bound estimates. As an example of the potential significance of large pteropod species, the ratio of large-pteropod mass flux at 400-m depths to the small-pteropod mass flux reported in the conservative account of Betzer *et al.* is about 5:1.

Analysis of our sediment trap collections in the northwest Pacific Ocean indicates that pteropod shells having diameters <500 μ m constitute at least 50% of the aragonite fluxes reported by Betzer *et al.*² for 100-m water depths. Shells in this size class, or smaller, should undergo extensive dissolution in the upper water column. Scanning electron microscopy of individuals in various pteropod size fractions provides direct evidence of extensive surface etching in the smaller size classes at relatively shallow depths. Figure 4 shows the contrasting dissolution features of the large rapid-settling pteropod species *C. columnella* and the small fragile species *L. inflata*. The smooth surface of *C. columnella*, even under considerable magnification (Fig. 4b), is typical of all large pteropod species collected in the deepest, 2,100 m, sediment traps. In contrast (Fig. 4c,d), *L. inflata* and other small pteropod species collected at 2,100-m depths exhibit extensive surface etching. At 2,100-m depths, slow-settling species have been in contact with undersaturated Pacific waters long enough (~1–2 days) to have lost a considerable portion of their original integument. The smallest size fraction in our collections gave evidence of advanced dissolution: the *L. inflata* shell (Fig. 5) has retained only a small fraction of its original smooth surface; Fig. 5a shows the aragonite surface which initially results from loss of this integument; Fig. 5b shows the considerable roughening of the aragonite surface on further dissolution. The shell transformations exhibited by the small specimen in this figure should be characteristic of a shell settling at about 1 cm s^{-1} at 35° N (Fig. 3a). The abundant small size class aragonitic particulates observed in our studies should provide substantial contributions to the alkalinity of the upper water column of the Pacific Ocean. As such, we think that pteropods should be identified as the principal source of 'excess alkalinity' in the Pacific Ocean, described by Fiadeiro³.

According to Berner Berner and Honjo²⁵ and Betzer *et al.*², aragonite appears to be a major component of the marine CO_2 system. The results presented here indicate that aragonitic particulates should undergo substantial transformations in the upper water column of the Pacific Ocean. The relatively rapid dissolution kinetics observed in this study and previously^{6,7} indicate that detailed experimental assessments of aragonite fluxes in the Pacific Ocean should be conducted in a manner minimizing sample transformations subsequent to collection. This work, in conjunction with the analysis of Betzer *et al.*, provides substantial evidence that the appearance of alkalinity anomalies, in waters substantially supersaturated with respect to calcite³ can be attributed to water column dissolution of aragonitic pteropods.

We thank Captain Raymond Speer and the crew of the RV *Discoverer* for their effective assistance during our analyses; chief scientist Joel Cline for administering diverse research activities at sea, Chief Bosun Warren Sherril for sediment trap deployments and retrievals; Dr. John A. McGowan for discussions on the characteristics of pteropods in the Pacific Ocean; Anthony Grew for scanning electron microscope analyses and Thomas Kuba for computer programming assistance. This research was supported by a grant (NA80RAD00020) from the National Oceanic and Atmospheric Administration Air

Resources Laboratory and is Contribution No. 740 from Pacific Marine Environmental Laboratory. We acknowledge a Graduate Council Fellowship award to J.G.A. from the University of South Florida.

Received 16 August; accepted 27 September 1984.

1. Berner, R. A. in *The Fate of Fossil Fuel CO₂ in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 243–260 (Plenum, New York, 1977).
2. Betzer, P. R. *et al. Science* (in the press).
3. Fiadairo, M. *Earth planet. Sci. Lett.* **49**, 499–505 (1980).
4. Broecker, W. S. in *Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 239–265 (Yale University Press, New Haven, 1971).
5. Broecker, W. S. in *The Fate of Fossil Fuel CO₂ in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 213–241 (Plenum, New York, 1977).
6. Morse, J. W., deKanel, J. & Harris, K. *Am. J. Sci.* **279**, 488–502 (1979).
7. Keir, R. C. *Geochim. cosmochim. Acta* **44**, 241–252 (1980).
8. Milliman, J. D. in *The Fate of Fossil Fuel CO₂ in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 641–654 (Plenum, New York, 1977).
9. Honjo, S. & Erez, J. *Earth planet. Sci. Lett.* **40**, 287–300 (1978).
10. Feely, R. A., Byrne, R. H., Betzer, P. R., Gendron, J. F. & Acker, J. G. *J. geophys. Res.* (in the press).
11. Millero, F. J. *Geochim. cosmochim. Acta* **43**, 1651–1661 (1979).
12. Turner, D. R., Whitfield, M. & Dickson, A. G. *Geochim. cosmochim. Acta* **45**, 855–881 (1981).
13. Morse, J. W. & Berner, R. A. in *Chemical Modeling in Aqueous Systems* (ed. Jenne, E. A.) 499–535 (American Chemical Society, Washington, D.C., 1979).
14. Mucci, A. *Am. J. Sci.* **283**, 780–799 (1983).
15. Berner, R. A. *Am. J. Sci.* **276**, 713–730 (1976).
16. Plath, D. C. thesis, Oregon State Univ. (1979).
17. MacIntyre, W. G. *Fish. Res. Bd. Canada Rep. Ser.* **200**, (1965).
18. Morse, J. W., Mucci, A. & Millero, F. J. *Geochim. cosmochim. Acta* **44**, 85–94 (1980).
19. Plummer, L. N. & Sundquist, E. T. *Geochim. cosmochim. Acta* **46**, 247–258 (1982).
20. Ingie, S. E., Culberson, C. H., Hawley, J. E. & Pytkowicz, R. M. *Mar. Chem.* **1**, 295–307 (1973).
21. Berger, W. H. *Science* **156**, 383–385 (1967).
22. Vinogradov, M. Ye. *Am. geophys. Un.* **136/141**, 39–42 (1961). (Transl. from Russian.)
23. Morse, J. W. cited in Berner, R. A. in *The Fate of Fossil Fuel CO₂ in the Oceans* (eds Andersen, N. R. & Malahoff, A.) 243–260 (Plenum, New York, 1977).
24. Adelseck, C. G. & Berger, W. H. *Cushman Foundation for Foraminiferal Research spec. Publ.* **13**, 70–81 (1975).
25. Berner, R. A. & Honjo, S. *Science* **211**, 940–942 (1981).

Characterization of the human factor VIII gene

Jane Gitschier, William I. Wood, Therese M. Goralka, Karen L. Wion, Ellson Y. Chen, Dennis H. Eaton, Gordon A. Vehar*, Daniel J. Capon & Richard M. Lawn

Departments of Molecular Biology and * Protein Biochemistry, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

The complete 186,000 base-pair (bp) human factor VIII gene has been isolated and consists of 26 exons ranging in size from 69 to 3,106 bp and introns as large as 32.4 kilobases (kb). Nine kb of mRNA and protein-coding DNA has been sequenced and the mRNA termini have been mapped. The relationship between internal duplications in factor VIII and evolution of the gene is discussed.

HAEMOPHILIA has been known for millennia to be a male-specific, inherited disease. Haemophilia A afflicts 10–20 per 100,000 males; a relatively high proportion due to novel mutations of the factor VIII gene^{2,3}. The high frequency of haemophilia A compared with other autosomal clotting disorders results from the location of this gene on the X chromosome; one affected gene will thus produce the disease state in males. We report here the isolation and characterization of the 186,000 base pair (bp) region of the human X chromosome containing the complete factor VIII gene. A set of overlapping bacteriophage and cosmid clones was used to map and sequence the 26 exons containing the approximately 9,000 bp of coding sequence. Accompanying articles describe the expression of active human factor VIII from recombinant DNA clones⁴ and present the detailed characterization of the protein⁵. Our results provide a basis for studies of the molecular defects associated with haemophilia.

Factor VIII genomic clones

Initial factor VIII clones were recovered from a bacteriophage λ genomic library⁴ (' λ 4X') containing DNA derived from an individual with 4X chromosomes (karyotyped 49,XXXXY). This library was screened with a unique 36-base oligonucleotide probe ('8.3'), synthesized to represent one codon choice of a sequenced tryptic peptide recovered from purified human factor VIII⁵. These initial clones contained overlapping segments spanning 28 kilobases (kb) of the human X chromosome. We subsequently expanded the set of genomic clones to contain 200 kb of the human genome encompassing the entire factor VIII gene. Both bacteriophage λ and cosmid clones were used in the process of 'genomic walking' (see legends to Figs. 1 and 2).

The restriction endonuclease map of the human factor VIII gene is shown in Fig. 1; the gene map was confirmed by the characterization of at least two independent clones for most of the 210-kb region and by numerous Southern blots of 49,XXXXY and normal DNA. No noticeable discrepancies between genomic DNA and the map derived from recombinant clones were found. In addition, low stringency blot experiments

have so far detected no closely related factor VIII genes or pseudogenes.

DNA sequence analysis

The DNA sequence of the entire mRNA coding portion (~9 kb) of the factor VIII gene has been determined. Exon-containing restriction fragments of genomic clones were identified by hybridization to cDNA and sequenced by dideoxy-chain termination procedures⁶. This DNA sequence analysis is distinct from that of cDNA clones derived from AL-7 cell RNA (see ref. 4). The genomic DNA sequence of exon boundaries (Fig. 3) and the promoter and polyadenylation regions (Fig. 4) are reported here. We found only two nucleotide differences between the genomic and cDNA sequences in the 8,860 nucleotides compared. Nucleotide 3,780 in cDNA is the G of a glutamic acid codon GAG, whereas in the cloned genomic DNA it is C, creating the conservative substitution of an aspartic acid residue. Nucleotide 8,728 in the 3' untranslated region is G in cDNA and A in genomic DNA. The surprising degree of similarity between genomic sequence derived from the 49,XXXXY cell DNA and cDNA sequence from AL-7 cell RNA indicates that the sequence data and the reverse transcription process are reliable and suggests also a lower level of nucleotide sequence polymorphism of this X-linked gene than has been found for some autosomal genes, an effect noted for other X chromosome loci⁷.

The boundaries of exons shown in Fig. 3 were established by restriction mapping and DNA sequence comparison of cDNA with genomic clones. In cases of nucleotide redundancy at either side of an intron, boundaries were chosen to be consistent with established consensus splice sites^{8,9}. All of the splice donor and acceptor sites conform to the GT...AG rule for nucleotides immediately flanking exon borders (Fig. 3). Further flanking sequences are in general agreement with favoured nucleotide frequencies noted in other compendia^{8,9}. The dinucleotide AG never occurs in the pyrimidine-rich region encompassing the final 15 nucleotides of introns; exons frequently begin with the sequence GT.

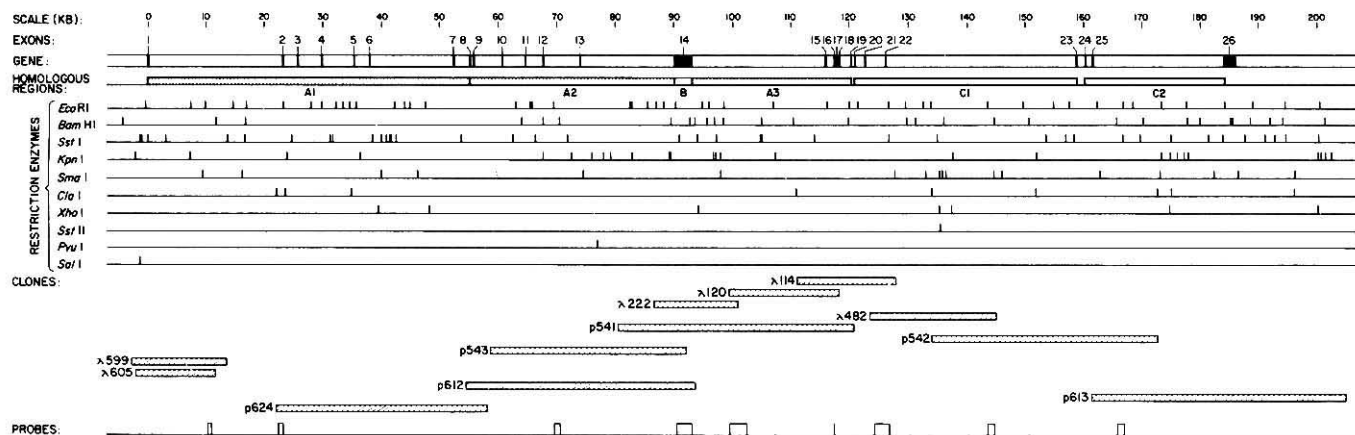
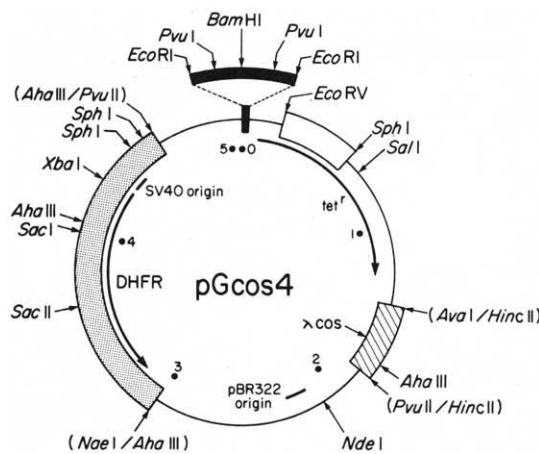


Fig. 1 Map of human factor VIII gene. The structure of the gene is schematically represented by an open bar; the 26 exons are filled-in areas drawn to scale. The direction of transcription is from left to right. The size scale in kb is drawn above the gene; the initiating ATG is positioned at nucleotide 1. Immediately below the gene are the extent of the triplicated 'A' domain, the unique 'B' domain, and the duplicated 'C' domain. The location of the recognition sites for the 10 restriction enzymes used to map the factor VIII gene are given in the next series of lines. Cross-hatched boxes represent the extent of human genomic DNA contained in each λ phage and cosmid clone. Other clones which encompass the same regions are not shown. No genomic clones span the 8.5 kb of intron located between clones λ 599 and p624. Other *EcoRI*, *BamHI* or *SstI* sites may lie in this region. The extent of this gap was determined by probing genomic Southern blots with single copy fragments isolated from the 3' end of λ 599 and the 5' end of p624. The bottom line shows the location of synthetic oligonucleotide or restriction fragment probes used in the genomic cloning, from left to right: the 2 probes used to map the uncloned region, three 5' walk probes, the original synthetic 36-mer probe '8.3'; three 3' walk probes.

Methods: λ and cosmid libraries were screened with random calf-thymus primed 32 P-labelled single copy genomic or cDNA probes²⁷. Nitrocellulose filters were hybridized overnight at 42 °C in 50% formamide, 0.1 g l⁻¹ salmon sperm DNA, 5 $\times$ SSC, 0.05 M sodium phosphate (pH 6.8), 5 $\times$ Denhardt's solution and 10% dextran sulphate²⁸. Filters were washed in 0.2 $\times$ SSC, 0.1 g l⁻¹ SDS at 60 °C for 1 h. Library filters were repeatedly rescreened with newly isolated walk probes so that extending clones were easily identified. To map restriction endonuclease cleavage sites, DNA from the clones was digested with restriction enzymes singly or in combinations and characterized by gel electrophoresis (followed by Southern blot hybridization in some cases). DNA fragments generated by *EcoRI* and *BamHI* digestion were subcloned into pUC plasmid vectors²⁹ for further propagation and analysis. Restriction mapping, DNA sequence analysis and blot hybridizations with the 8.3 probe determined the orientation of transcription of the gene. The genome walk was initiated by identifying single copy probe fragments near the ends of the initial 28 kb region. To find suitable fragments, digests of cloned DNA were blot hybridized with total 32 P-labelled human DNA under conditions where only DNA fragments containing sequences repeated more than about 50 times in the genome will hybridize^{30,31}. Candidate fragments were retested for repeated sequences by hybridization to 50,000 phage from the λ /4X library. To begin extension in the 5' direction, clone λ 222 was identified by a triplet of 1 kb *NdeI*/*BamHI* probe fragments isolated from λ 120 DNA. In the 3' direction, a single copy probe fragment of λ 114 failed to yield extending clones in screens of several λ libraries. Based on genomic blotting results, we then constructed a size selected³² *BclI* library of human 49,XXXXY DNA in the vector λ 1059³³. λ 482 is one of the extending clones so derived. Subsequent overlapping clones (p541, p542, p543, p612, p613 and p624) were derived by screening a recombinant cosmid library containing human DNA (see Fig. 2). The most 5' genomic clones, λ 599 and λ 605 were obtained by screening with cDNA-derived probes.

Fig. 2 Cosmid vector pGcos4. To facilitate genomic walking, we developed a new cosmid vector with the following features: (1) It can accommodate 45 kb inserts. (2) It confers resistance to tetracycline, rather than ampicillin. (3) The 641 bp *AvaI*/*PvuII* fragment of pBR322 has been removed to increase plasmid copy number and to delete sequences interfering with transformation of eukaryotic cells³⁴. (4) A methotrexate-resistant dihydrofolate reductase gene driven by a simian virus 40 (SV40) early promoter has been incorporated for the selection and propagation of clones in eukaryotic cells³⁵. (5) A synthetic DNA fragment containing a unique *BamHI* site flanked by pairs of *PvuI* and *EcoRI* sites has been inserted as the cloning site. The flanking *EcoRI* sites are useful for subcloning fragments and the *PvuI* sites allow excision of the entire insert in most cases, since this enzyme cuts eukaryotic DNA only about once per 135,000 bp. The scale is represented by dots within the circle numbered in kb from the zero point. The 403 base (b) annealed *HincII* fragment of λ c1857S7 (Bethesda Research Lab) containing the *cos* site was cloned in pBR322 from *AvaI* to *PvuII* to generate the plasmid pGcos1. Separately, the 1,624 b *PvuII* to *NaeI* fragment of pFR400³⁵ containing an SV40 origin and promoter, a methotrexate-resistant dihydrofolate reductase gene, and hepatitis B surface antigen polyadenylation sequences, was cloned into the pBR322 *AhaIII* site to generate the plasmid mp33dhfr. The following three fragments were ligated together to generate the cosmid vector pGcos3: 1,497 base (b) *SphI*-*NdeI* fragment of pGcos1, 2,163 b *NdeI*-*EcoRV* fragment of mp33dhfr and 376 b *EcoRV*-*SphI* fragment of pKTI9. pKTI9 is a derivative of pBR322 in which the *BamHI* site in the tetracycline resistance gene has been removed (provided by Herb Boyer). pGcos4 was generated by cloning the synthetic 20mer, 5' AATTCGATCGGATCCGATCG, in the *EcoRI* of pGcos3. Left and right arms of this vector were prepared by cleavage of two aliquots with *SstI* or *SalI*, followed by treatment with alkaline phosphatase, cleavage with *BamHI* and isolation of the large fragments²⁸. 49,XXXXY DNA isolated from the GM1202A cell line (NIGMS Human Genetic Mutant Cell Repository, Camden, New Jersey) was partially cleaved with five levels of *Sau3A1* and the 45 kb fragments isolated by sucrose density centrifugation²⁸. These fragments were ligated to the *BamHI* arms of pGcos4, packaged *in vitro* and used to infect *E. coli* HB101.



©1984 Nature Publishing Group

Fig. 5 Determination of the 5' mRNA start site by RNase protection. Autoradiograph of a 6% polyacrylamide, 7 M urea gel. Lanes 1 and 2, a known DNA sequencing ladder employed as size standards; lanes 3–11, results of RNase protection reactions.

Methods: Cell and tissue RNA was prepared³⁶. The 5' genomic *Sst*I fragment was ligated into pSP64 (a gift of P. Mellon and T. Maniatis) and linearized with *Pvu*II. The uniformly labelled ³²P-RNA probe was synthesized with 400 Ci mmole⁻¹ α -[³²P]CTP (Amersham) and SP6 RNA polymerase (NEN) followed by DNase treatment (DPRF, Worthington) as described^{12,13}. Sample RNAs were hybridized with 5×10^5 c.p.m. of probe at 37°C and digested with 80 μ g ml⁻¹ RNase A and 4 μ g ml⁻¹ RNase T1 (Worthington) followed by extraction, electrophoresis and autoradiography. ³²P-labelled, anti-sense strand RNA probe was prepared by ligating the 1.2 kb *Sst*I genomic fragment containing exon 1 and 5' flanking sequences into pSP64¹² and transcribing the linearized plasmid with SP6 RNA polymerase^{12,13}. This labelled probe was hybridized with poly(A)⁺ RNA derived from various sources and digested with RNase A and RNase T1 before electrophoresis. Protected RNA was derived from: lane 3, AL-7 human hybridoma cells (10 μ g); 4, human liver sample 1 (10 μ g); 5, human liver sample 2 (10 μ g); 6, human Alexander cell line (10 μ g); 7, human Hep G2 cell line (10 μ g); 8, peripheral blood lymphocyte sample 1 (3 μ g); 9, peripheral blood lymphocyte sample 2 (3 μ g); 10, mouse T lymphoma cell line S49 (10 μ g); 11, no RNA control. An identical pair of protected fragments (arrow), measuring 185 and 187 bases from the *Sst*I site, appear only in RNA obtained from the two human liver samples and the AL-7 hybridoma cells. This positions the probable 5' mRNA start site at ~170 bases 5' of the initiation ATG as shown in Fig. 4. (As RNases A and T1 cannot cleave after an adenosine residue and the complementary strand is radiolabelled, no protected band was seen corresponding to a terminus at position -171 of the gene (Fig. 4), which is a T residue in the genomic sequence.)

the protected RNA is bounded by an intron instead of the true mRNA 5' start site.

The RNase protection experiments also show that liver is a site of synthesis of factor VIII. No factor VIII RNA was detected in the other tissues of the protection experiment shown in Fig. 5 nor by Northern blot analysis of RNA from various tissues and cell lines⁴. Liver is only one of several candidates for the site of factor VIII production^{14–16}.

Inspection of the genomic sequence of the 5' flanking region (Fig. 4a) revealed the sequence GATAAA beginning 30 bp 5' of the presumed 5' mRNA start site, similar to the canonical 'ATA' or 'Goldberg-Hogness' sequence, located about 25–35 bp before eukaryotic cap sites and apparently required for precise initiation of transcription by RNA polymerase II^{17,18}. We found no convincing candidate for the less frequently conserved 'CAT' sequence¹⁹ in the -70 bp region.

One additional ATG to nucleotide 5' to +1 is found in the 5' untranslated region at position -131 (Fig. 5). This region cannot serve as the translation start site for factor VIII because seven stop codons, distributed in all three possible reading frames, occur in the 128 nucleotides between this upstream ATG and position +1. Although only 5% of the previously compiled eukaryotic mRNAs contain an ATG 5' of the translation initiation site²⁰, the presence of an upstream ATG does not necessarily repress translational initiation at the proper start site when it is followed by an in-frame termination codon²¹.

The 3' untranslated region of the factor VIII gene contains 1,805 nucleotides (counting the TGA stop triplet and terminating before the first A of the poly(A) tail; Fig. 4b). The site of polyadenylation was inferred from cDNA clones and is preceded 19 bases by the common polyadenylation signal AATAAA²² as well as a second, less conserved recognition element CATTG²³.

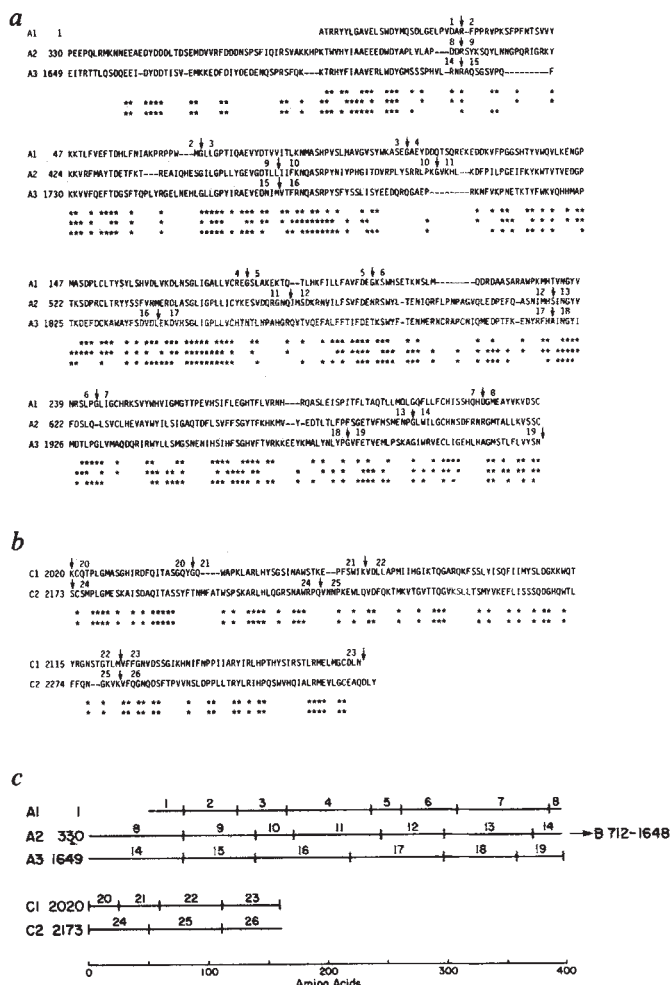


Fig. 6 Homologous factor VIII domains and intron boundaries.

a, 'A' unit triplication showing the alignment of three amino acid sequences. The amino acid numbers precede each line. Arrows indicate the location of introns and are flanked by the appropriate exon numbers. *, Conserved amino acids. **b**, 'C' unit duplication. **c**, Representation of the A and C repeat units with the location of introns indicated by vertical lines. The exons are numbered.

Introns and protein domains

Searches for internal homologies in the factor VIII DNA sequence and the derived amino acid sequence revealed three types of domains: a triplicated 'A' unit of 330–380 amino acids, a unique 'B' unit of about 925 amino acids and a duplicated 'C' unit of about 160 amino acids. The order of the domains in the gene and protein is A1–A2–B–A3–C1–C2 (refs 4, 5; Fig. 1) (note that the exons constituting each domain are not clustered). The amino acid alignment of the A and C repeats (Fig. 6) demonstrates that the A units have ~30% pairwise amino acid homology (20% of the residues being conserved in all three domains), whereas the two C units have 37% homology^{4,5}.

If segmental DNA duplication has played a role in the evolution of the large factor VIII gene, one might expect a conservation of the intron boundaries within the A and C repeats of factor VIII. The alignment of most intron boundaries relative to the protein coding sequence is well fixed by the high degree of homology in the A and C repeats (Fig. 6). For the C duplication, intron boundaries occur precisely at the borders of the C1 and C2 repeats, as would be expected for a DNA duplication mechanism where the boundaries fall within the introns. Furthermore, the junction between A3 and C1 also contains an intron precisely at the boundary, again as expected for an intron joining mechanism. On the other hand, the A1–A2 and A2–A3 junctions are each contained on one exon (Fig. 6).

Within the A and C repeats only some of the intron boundaries

are conserved. In fact, each of the repeats contains a different number of exons. The precise alignment of some intron boundaries within the A and C repeats suggests that these introns were present in an ancestral precursor of the repeat units and duplicated in the course of the event that generated the repeat. The non-aligned introns appear to have arisen since the duplication, either by loss of one member of a duplicated intron or by the insertion of new introns after the duplication. In an analogous situation, α -feto protein and albumin each have a triplicated gene structure, but unlike factor VIII, all of the nine intron boundaries within the repeats for these two genes are conserved precisely²⁴⁻²⁵.

It is also interesting to speculate as to the origin of domain B in factor VIII. The 925 amino acids of this domain encompass nearly all of the 3.1 kb of exon 14, where the end of the A2 repeat (at the 5' end) and the start of the A3 repeat (at the 3' end) are also found. This suggests that during the course of evolution a processed gene may have been inserted into a short exon that contained the A2-A3 boundary. Insertion of a processed gene (derived from an mRNA) would also account for the anomalously large size of exon 14. The B domain connects the M_r 90,000 (made up of A1 and A2) and the M_r 80,000 (A3, C1, C2) fragments of the factor VIII protein. Proteolytic removal of the domain is associated with activation of the protein^{5,11};

thus exon 14 is related to a physiological unit of the protein. Whether any other functional domains or subdomains of factor VIII correspond to any of the exons awaits further characterization of the molecule.

In conclusion, we report here the complete isolation and characterization of the human factor VIII gene. It is the largest gene yet characterized, encompassing nearly 0.1% (186 kb) of the human X chromosome. Transcription of such a large gene could take over three hours (based on a transcription rate of 15 nucleotides per second²⁶). Probes derived from the factor VIII gene will allow identification of restriction fragment length polymorphisms which can be used further to localize the gene on the X chromosome and lead to improved prenatal diagnosis of haemophilia. We anticipate that the knowledge of the gene structure and DNA sequence will be invaluable in the analysis of the molecular basis of haemophilia.

We thank John McGrath who characterized the λ 222 clone and Philip Hollingshead for laboratory assistance; Edward Tuddenham, David Martin, Ingrid Caras and Glenn Nedwin for providing tissue samples; Jeanne Arch and Alane Gray for preparation of the manuscript and figures. We thank David Goeddel, Christian Simonsen, Laurence Lasky and Dennis Kleid for helpful discussions and David Martin and Robert Swanson for their support.

Received 9 August; accepted 27 September 1984.

1. Rosner, F. *Medicine in the Bible and the Talmud* 43-49 (Yeshiva University Press, New York, 1977).
2. Galjaard, H. *Genetic Metabolic Disease* (Elsevier, New York, 1980).
3. Haldane, J. B. S. *J. Genet.* **31**, 317-326 (1935).
4. Wood, W. I. *et al. Nature* **312**, 330-337 (1984).
5. Vehar, G. A. *et al. Nature* **312**, 337-342 (1984).
6. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
7. Cooper, D. N. & Schmidke, J. *Hum. Genet.* **66**, 1-66 (1984).
8. Breathnach, R. & Chambon, P. A. *Rev. Biochem.* **50**, 349-383 (1981).
9. Mount, S. M. *Nucleic Acids Res.* **10**, 459-472 (1982).
10. Blake, C. *Nature* **308**, 535-537 (1983).
11. Fulcher, C. A., Roberts, J. R. & Zimmerman, T. S. *Blood* **61**, 807-811 (1983).
12. Melton, D. A. *et al. Nucleic Acids Res.* (in the press).
13. Zinn, K., Di Maio, D. & Maniatis, T. *Cell* **34**, 865-879 (1983).
14. Zimmerman, T. S. & Meyer, D. in *Haemostasis and Thrombosis* (eds Bloom, A. L. & Duncan, P. T.) 111-123 (Churchill Livingstone, New York, 1981).
15. Stel, H. V., van der Kwast, T. H. & Veerman, E. C. I. *Nature* **303**, 530-532 (1983).
16. Kelly, D. A., Summerfield, J. A. & Tuddenham, E. G. D. *Br. J. Haem.* **56**, 535-543 (1984).
17. Corden, J. *et al. Science* **209**, 1406-1414 (1980).

18. Grosschedl, R. & Birnstiel, M. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1432-1436 (1980).
19. Benoist, C. *et al. Nucleic Acids Res.* **8**, 127-142 (1980).
20. Kozak, M. *Nucleic Acids Res.* **8**, 127-142 (1980).
21. Liu, C.-C., Simonsen, C. C. & Levinson, A. D. *Nature* **309**, 82-85 (1984).
22. Proudfoot, N. & Brownlee, G. *Nature* **252**, 359-362 (1981).
23. Berget, S. M. *Nature* **309**, 179-181 (1984).
24. Eiferman, F. A. *et al. Nature* **294**, 713-718 (1981).
25. Sargent, T. D. *et al. Molec. Cell Biol.* **1**, 871-883 (1981).
26. Davidson, E. H. *Gene Activity in Early Development*, 2nd edn 363 (Academic, New York, 1976).
27. Taylor, J. M., Illmensee, R. & Summer, S. *Biochim. biophys. Acta* **442**, 324-330 (1976).
28. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
29. Vieira, J. & Messing, J. *Gene* **19**, 259-268 (1982).
30. Wood, W. I., Nichol, J. & Felsenfeld, G. *J. biol. Chem.* **256**, 1502-1505 (1981).
31. Shen, C.-K. J. & Maniatis, T. *Cell* **19**, 379-391 (1981).
32. Lawn, R. M. *et al. Nucleic Acids Res.* **9**, 6103-6114 (1981).
33. Korn, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 5172-5176 (1980).
34. Lusky, M. & Botchan, M. *Nature* **293**, 79-81 (1981).
35. Simonsen, C. C. & Levinson, A. D. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2495-2499 (1983).
36. Ullrich, A. *et al. Science* **196**, 1313-1315 (1977).

Expression of active human factor VIII from recombinant DNA clones

William I. Wood, Daniel J. Capon, Christian C. Simonsen, Dan L. Eaton*, Jane Gitschier, Bruce Keyt*, Peter H. Seeburg, Douglas H. Smith, Philip Hollingshead, Karen L. Wion, Eric Delwart, Edward G. D. Tuddenham†, Gordon A. Vehar* & Richard M. Lawn

Departments of Molecular Biology and *Protein Biochemistry, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

† Haemophilia Centre, Royal Free Hospital School of Medicine, London WC1N 1BP, UK

DNA clones encoding the complete 2,351 amino acid sequence for human factor VIII have been isolated and used to produce biologically active factor VIII in cultured mammalian cells. The recombinant protein corrects the clotting time of plasma from haemophiliacs and has many of the biochemical and immunological characteristics of serum-derived factor VIII.

A CRITICAL factor in the arrest of bleeding is the activation of the coagulation cascade, a system consisting of over a dozen interacting proteins present in plasma or released from cells^{1,2}. A small stimulus at the beginning of the cascade is catalytically amplified at each step such that the final outcome is the rapid conversion of the soluble protein fibrinogen into its insoluble

form, fibrin, which stabilizes the platelet plug. The lack or deficiency of any of the proteins involved in the coagulation cascade blocks the propagation of the initial stimulus. In the middle of the cascade is a step wherein factor IXa, in conjunction with factor VIII, converts factor X to the activated form, factor Xa. Factor VIII acts as a cofactor at this step, being required

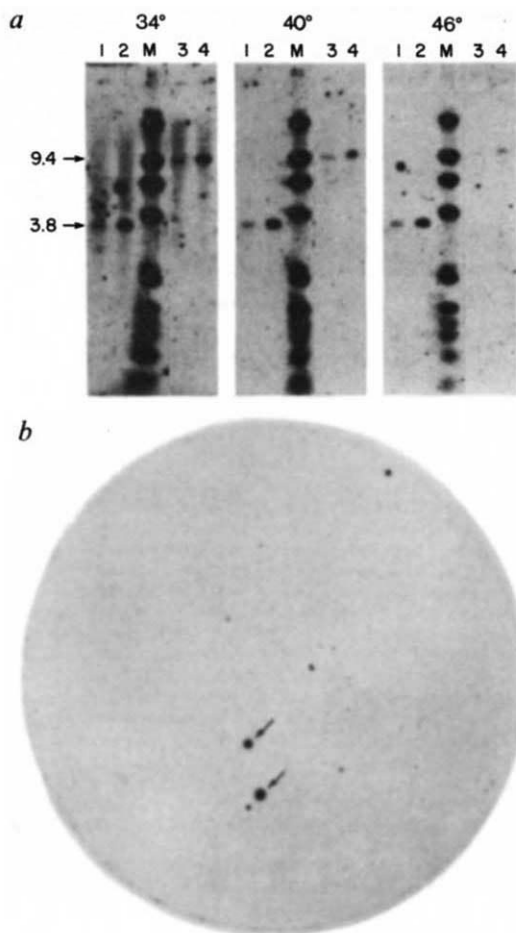


Fig. 1 Detection of the factor VIII gene with probe 8.3, 5'-CTTTTCCAGGTCAACGTCGGAGAAATAAGCCCAAGC, a 36-base long probe synthesized based on the sequence of a 12 amino acid tryptic fragment of factor VIII, AWAYFSDVDLEK (ref. 5). **a**, Southern blots of 46,XY (1X, male) DNA and 49,XXXXY (4X) human DNA (isolated from cell line GM1202A, NIGMS Human Genetic Mutant Cell Repository, Camden, New Jersey) digested with *EcoRI* and *BamHI* were hybridized with ^{32}P end-labelled 8.3 probe in $6\times\text{SSC}$, 50 mM sodium phosphate (pH 6.8), $5\times\text{Denhardt's}$ solution, 0.1 g l^{-1} boiled, sonicated salmon sperm DNA, 20% formamide and 10% dextran sulphate at 42°C . The three blots were washed in $1\times\text{SSC}$, 0.1% SDS at the temperature indicated. Lane 1, *EcoRI* 1X; lane 2, *EcoRI* 4X; lane 3, *BamHI* 1X; lane 4, *BamHI* 4X; lane M, end-labelled λ HindIII and ΦX174 *HaeIII* digested marker fragments. **b**, One nitrocellulose filter from the screen of cloned 4X human DNA hybridized with probe 8.3. This library was constructed by partially digesting 4X DNA with *Sau3AI* and ligating the 15 kb size fraction into the vector λ Charon 30 (ref. 42). Arrows, the independent factor VIII positive clones. Hybridization and washing for the library screen was as described above for the Southern blots with a wash temperature of 37°C .

with calcium ions and phospholipid for the activity of factor IXa. The two most common haemophilic disorders are caused by the decreased function of either factor VIII (haemophilia A or classic haemophilia; 80% of all cases) or factor IXa (haemophilia B or Christmas disease). The clinical manifestation in both types of disorders is the same: a lack of sufficient fibrin formation resulting in prolonged bleeding.

Several decades ago the mean age of death of haemophiliacs was less than 20 years; between 1950 and 1970, however, research into the factor VIII disorder led to the treatment of haemophilia A, first with whole plasma and, later, with plasma concentrates. Human plasma is both an inefficient and expensive source of this labile protein which is present at only $100\text{--}200\text{ ng ml}^{-1}$ (ref. 3). Pharmaceutical products currently available are highly impure ($<1\%$ factor VIII) and some are produced

Table 1 Assays of recombinant factor VIII activity

Sample	Coatest assay (units ml^{-1})	Radioimmune assay (units ml^{-1})	
		C7F7 Ab	C10 Ab
62.2 supernatant	0.07	0.28	0.34
62.2 supernatant + C7F7	<0.01	—	—
62.2 supernatant + Synbiotic	<0.01	—	—
Control supernatant	<0.01	<0.001	<0.001
62.2 cell lysate	—	0.1	0.1

Culture supernatant from recombinant factor VIII producing 62.2 cells, with and without inhibiting monoclonal antibodies, was subjected to the chromogenic 'Coatest' or to radioimmune assays. The Coatest assay (Helena Labs) was performed by diluting $5\text{ }\mu\text{l}$ of 62.2 supernatant with $45\text{ }\mu\text{l}$ 0.05 M Tris, pH 7.3 containing 0.2% bovine serum albumin (BSA) and incubating this mixture with or without either $5\text{ }\mu\text{g}$ C7F7 or Synbiotic factor VIII monoclonal antibodies. Both these antibodies react with the $M_r 80,000$ subunit of plasma-derived factor VIII. After 5 min at 24°C , $50\text{ }\mu\text{l}$ of a mixture containing factor IXa, factor X and phospholipid was added and samples incubated for 5 minutes at 37°C . $50\text{ }\mu\text{l}$ of each of 25 mM CaCl_2 and S-2222 chromogenic substrate was added and after 10 min at 37°C the reaction stopped by the addition of 50% acetic acid. The assay was repeated with control cultures and calibrated with normal human plasma. One factor VIII unit is equivalent to the amount of factor VIII in 1 ml normal human plasma. Radioimmune assays: The 96 wells of a microtitre dish were coated overnight with $100\text{ }\mu\text{l}$ 50 mM NaHCO_3 buffer, pH 9.6, containing 2.5 mg l^{-1} of C.C. antibody purified by protein A-Sepharose chromatography. The wells were washed three times with $200\text{ }\mu\text{l}$ of phosphate-buffered saline (PBS) containing 0.05% Tween 20, blocked with $200\text{ }\mu\text{l}$ of PBS containing 0.1% gelatin and 0.01% merthiolate for 1–2 h, rewashed as before and $100\text{ }\mu\text{l}$ of sample added and incubated overnight. The wells were washed and $100\text{ }\mu\text{l}$ of ^{125}I labelled C10 or C7F7 antibody ($1,000\text{ c.p.m. }\mu\text{l}^{-1}$) added and incubated for 6–8 h. The wells were washed again and counted. Standard curves were derived from samples of normal plasma. Antibodies used in assays and columns were: C.C., a polyclonal antibody derived from the plasma of a severely affected haemophilic⁴⁰; C8 and C10, neutralizing monoclonals binding to the $M_r 210,000$ and $90,000$ fragments of factor VIII, supplied by A. H. Goodall, Department of Immunology, Royal Free Hospital, London⁴¹; C7F7, a neutralizing monoclonal binding to the $M_r 80,000$ fragment of factor VIII, derived at Genentech by Badora Streck and an $80,000$ binding monoclonal obtained from Synbiotic Corporation.

from pooled plasma lots derived from thousands of donors. Thus, these products are associated with a variety of serious complications caused by protein precipitates as well as the possible contamination by adventitious agents such as hepatitis virus and the agent responsible for acquired immune deficiency syndrome (AIDS).

Factor VIII is a molecule of relative mass (M_r) 300,000 (ref. 4). In attempts to produce pure homogenous factor VIII, we have identified clones containing the factor VIII gene using oligonucleotide probes based on the protein sequence data reported in an accompanying paper⁵. We also show that cloned full-length factor VIII protein coding sequences assembled into a plasmid direct the synthesis of biologically active factor VIII when transfected into cultured mammalian cells.

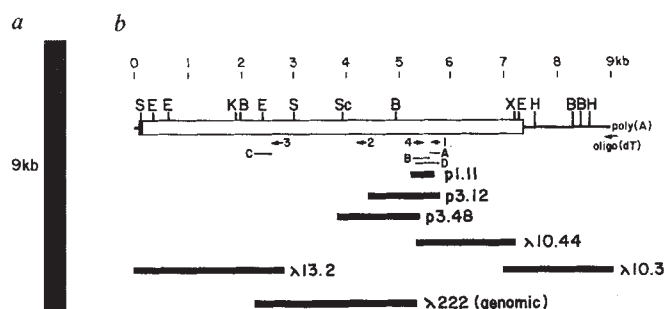
Isolation of genomic clones

Two factors contributed to our decision to initially screen recombinant genomic libraries (rather than RNA-derived cDNA libraries) for the factor VIII gene. First, the site of synthesis of factor VIII is uncertain, although the liver is considered the most likely source^{6–10}. Second, the concentration of factor VIII in plasma is extremely low (200 ng ml^{-1}), about $1/2,000,000$ the molar concentration of serum albumin. Thus, it was not clear that cDNA libraries made from mRNA of a given tissue would yield any factor VIII clones.

Therefore we adopted a strategy of first identifying a genomic clone corresponding to a sequenced portion of the factor VIII protein. Then several paths could be followed. An exon-containing restriction fragment could be used to identify a

Fig. 2 cDNA cloning procedure. **a**, 25 μ g poly(A)⁺ RNA from the AL-7 cell line was electrophoresed in a 1% agarose gel containing 6% formaldehyde⁴², transferred to nitrocellulose and hybridized in 5 $\times$ SSC, 5 $\times$ Denhardt's solution, 10% dextran sulphate, 50% formamide, 0.1% SDS, 0.1% sodium pyrophosphate, 0.1 mg ml⁻¹ *E. coli* tRNA at 42°C overnight with ³²P-labelled probe prepared from the 189 bp *Stu*I/*Hinc*II genomic clone fragment containing exon 16 sequence. Wash conditions were 0.2 $\times$ SSC, 0.1% SDS at 42°C. This RNA was 0.001–0.001% of the total cytoplasmic poly(A)⁺ RNA in the AL-7 cell line by comparison with control DNA dot-blot hybridizations⁴³. **b**, Factor VIII mRNA; open bar, mature protein coding region; hatched area, signal peptide coding region; adjacent lines, untranslated regions of the message. Left, the 5' end of the mRNA. Below the mRNA line are the six cDNA clones and the genomic clone containing exon 14 from which the full length factor VIII clone was assembled. cDNA synthesis primers 1–4 and oligo(dT) are shown with arrows depicting the direction of synthesis for which they are primed. Primer 1, 5' CAGGTCAA-CATCAGAG; primer 2, 5' TATCTGTGTGAGGGTG; primer 3, 5' AACTCTGTTGCTGCAG; primer 4, 5' TATTGCTGCAGTGGAG. Probe A, 189 bp *Stu*I/*Hinc*II fragment from exon 16; probe B, 361 bp *Sau*3A1/*Stu*I fragment of p1.11; probe C, 324 bp *Hinf*I fragment from exon 14; probe D, 419 bp *Pst*I/*Hinc*II from p1.11. The original 36-mer probe 8.3 is located near the primer 1 region. Restriction endonuclease sites in the corresponding DNA are indicated by: B, *Bam*HI; E, *Eco*RI; H, *Hpa*I; K, *Kpn*I; S, *Sac*I; Sc, *Sca*I; X, *Xba*I.

Methods: the first cDNA clones were isolated by priming cDNA synthesis with primer 1 from exon 16. From 5 μ g AL-7 poly(A)⁺ RNA, approximately 100,000 *E. coli* transformants were obtained by introducing the double-stranded cDNA into pBR322 by the oligo(dC):dG tailing method⁴². These clones were screened by hybridization to probe A⁴⁴ and a single clone containing a 447 bp cDNA insert (p1.11) (including the 5' 104 bp of exon 16) was isolated. The remaining sequences of p1.11 extended 5', including exon 15 and part of exon 14. Exon 14 had been found earlier by hybridization with an oligonucleotide probe based on protein sequence data¹⁶. Hybrid selection was used to further enrich for factor VIII specific sequences in AL-7 mRNA. To avoid degradation of the relatively large factor VIII transcripts, single-stranded cDNA was initially synthesized on a preparative scale and subjected to hybrid selection prior to conversion to double-stranded molecules. First-strand synthesis was performed on 200 μ g AL-7 poly(A)⁺ RNA with primer 1 and the resulting product was hybridized with an exon 16 DNA fragment (probe A), denatured and immobilized on DBM-cellulose paper⁴². After elution, 0.5–1.0 ng of single-stranded material was recovered and used to prepare 12,000 *E. coli* transformants. 29 clones were obtained by hybridization to probe B which did not overlap the DNA fragment used for hybrid selection. This represents a 240-fold enrichment of desired clones compared with the first cDNA library. DNA sequence shows that the two longest clones (p3.12, p3.48) extended about 1,500 bp further 5' than p1.11. Primer 3, located near the 5' end of exon 14, was used to isolate the N-terminal coding sequences for factor VIII. Double stranded cDNA was prepared and ligated to synthetic complementary *Eco*RI adaptors, 5'pCCTTGACCGTAAGACATG and 5'AATTCATGTCTTACGGTCAAGG. Upon annealing, these adaptors have a phosphorylated blunt end which will be covalently joined to the cDNA, as well as a cohesive *Eco*RI terminus with a 5'-hydroxyl group which cannot self-ligate, thus avoiding the need for methylation of the cDNA and subsequent *Eco*RI digestion. Approximately 1.5 $\times$ 10⁶ recombinant phage were obtained from 1 μ g AL-7 poly(A)⁺ RNA with primer 3 following removal of excess adaptor, insertion into λ gt10 and *in vitro* packaging^{42,45}. Forty-six clones were identified by hybridization with probe C. Several of the clones contained cDNA inserts which extended about 2,700 bp 5' of primer 3 and the DNA sequence of the two longest clones (λ 13.2 and λ 13.27) revealed the 5' terminus of the protein coding region (Fig. 3). (Primer 2 was only used subsequent to the cloning and expression described here to obtain cDNA sequence data corresponding to the region of genomic clone λ 222 containing exon 14.) To isolate the coding region for the C-terminus of factor VIII, a cDNA library was prepared from oligo(dT) primed AL-7 RNA. A 16-mer corresponding to the mRNA sequence found 400 bp upstream of the 3' end of exon 16 (primer 4) was included in the second strand reaction to increase its specificity. Four of a total 3 $\times$ 10⁶ λ gt10 clones hybridized with a fragment from p3.12 (probe D) located 3' of primer 4. DNA sequence analysis of the longest clone, λ 10.44, indicated that the cDNA insert began at primer 4 and extended for nearly 1,800 bp 3', but had not reached the end of the coding region. To identify clones containing the 3' end of the mRNA, the oligo(dT)-primed library was rehybridized with a probe prepared from the insert of λ 10.44. Twenty-four additional clones were recovered and the two longest of these (λ 10.3 and λ 10.9) were sequenced. Both clones overlapped λ 10.44 as expected and extended an additional 1,900 bp to the poly(A) tail.



suitable tissue or cell source of factor VIII by RNA blot hybridization. The mRNA from this source would be used to generate cDNA clones which could be screened by cloned genomic fragments. Concurrently, overlapping genomic clones could be isolated encompassing the entire factor VIII gene. Even if a direct mRNA source were not identified, exons from the genomic clones could be found and spliced together by simian virus 40 (SV40) recombinant 'exon expression' plasmids. In fact, we used information derived from all three paths.

A genomic library enriched for the X chromosome was constructed from DNA isolated from a lymphoblast cell line derived from an individual with four X chromosomes (karyotyped 49,XXXXY). To identify factor VIII clones, a series of oligonucleotide probes based on portions of the human factor VIII protein sequence⁵ was prepared by chemical synthesis. The oligonucleotide probes were of two types: unique long probes of 30–100 nucleotides based on codon usage analysis and mixtures of short probes of 14–20 bases specifying all possible combinations for each codon choice^{11–13}. The utility of probe 8.3 (Fig. 1 legend) was first tested by genomic blot hybridizations with normal male (1X) and 49,XXXXY (4X) DNA. Even at the highest stringency (1 $\times$ SSC, 46°C), a single *Eco*RI or *Bam*HI band was observed which had an intensity ratio of about 1:4 in the 1X and 4X lanes, as expected for the X-linked factor VIII gene (Fig. 1).

Next, 500,000 phage of the 4X genomic library were screened with probe 8.3 and 15 clones representing eight distinct overlapping recombinants were isolated and characterized (Fig. 1). The DNA sequence of the region hybridizing with the 8.3 probe included two continuous matches of 14 and 10 base pairs (bp) and had an overall homology of 83%. The first 10 of the 12 amino acid residues of the peptide fragment agreed with those deduced from the DNA sequence of the recombinant clones. The DNA sequence also predicted a lysine residue preceding

the peptide, as would be expected for the product of a tryptic digest. The final two residues did not match the DNA-predicted sequence; the DNA at this junction, however, contained a probable RNA splice donor sequence^{14,15} followed shortly by stop codons in all three possible reading frames.

Inspection of the open reading frame immediately 5' of the probe homology revealed matches with the protein sequence of several additional tryptic peptides from the *M*_r 80,000 factor VIII fragment⁵. The identity between highly purified factor VIII

Fig. 3 Sequence of human factor VIII gene, with nucleotides numbered at the left of each line. One, the A of the translation initiation codon ATG. Negative numbers, 5' untranslated sequence. (mRNA mapping experiments suggest that factor VIII mRNA extends ~60 nucleotides further 5' than position -109 shown here¹⁶.) The predicted protein sequence is shown above the DNA. Numbers above the amino acids are S1–19 for the predicted signal peptide, and 1–2,332 for the predicted mature protein. 'Op', the opal translation stop codon TGA. The 3' polyadenylation signal AATAAA¹⁸ is underlined and eight residues of the poly(A) tail (found in clone λ c10.3) are shown. The sequence homologous to the synthetic oligonucleotide probe 8.3 is also underlined (nucleotides 5,557–5,592). Selected restriction endonuclease cleavage sites are shown above the appropriate sequence. The entire factor VIII sequence was verified after assembly of cDNA and genomic clones in an expression plasmid. The complete DNA sequence of the protein coding region of the human factor VIII gene was also determined from the genomic clones we have described¹⁶. cDNA clones have been isolated covering all of the sequence shown except for nucleotides 3,042–3,216. The cDNA and genomic clones differ at only two positions: nucleotide 3,780 (underlined) is C in the genomic clone and G in cDNA (changing amino acid 1,241 from Glu to Asp), and nucleotide 8,728 (underlined) in the 3' untranslated region is A in the genomic clone and G in cDNA. DNA sequence analysis was performed by the dideoxy chain termination method using restriction fragments subcloned into single-stranded M13 vectors^{46,47}.

[illegible][illegible]

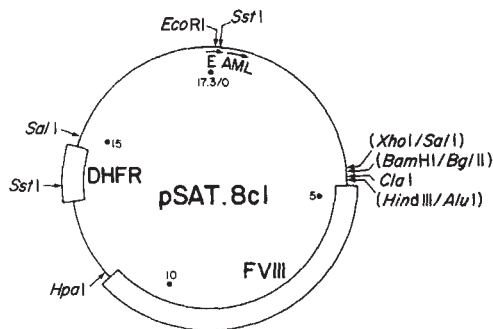


Fig. 4 Assembly of full-length factor VIII expression plasmid. Restriction sites delimiting various DNA fragments that make up pSAT.8c1 (initially referred to as pAML3P.8c1) are shown. Not all the sites for a particular enzyme are shown. The scale in kb is shown on the inside by dots. The factor VIII (FVIII) and dihydrofolate reductase (DHFR) gene coding regions are shown as well as the tandem SV40 early (E)/adenovirus major late (AML) promoter. Beginning with the *SalI* near 15 kb and proceeding clockwise, the various pieces of the plasmid are: *SalI* to *EcoRI*, p342E⁴⁸ fragment containing the pBR322 origin and ampicillin resistance gene and an SV40 early promoter; *EcoRI* to *SstI*, linker of pUC13⁴⁶; *SstI* to *XhoI*, adenovirus-2 (Bethesda Research Laboratories) fragment coordinates 15.4 to 26.5⁴⁹ containing the major late promoter; *SalI* to *BamHI*, linker of pUC13; *BglII* to *ClaI*, short linker with the sequence 5' GATCTCCTAGACACCGCTCAGCTCTGTAT; *ClaI* to *HindIII* short linker from pBR322 with the *HindIII* site filled in with Klenow polymerase; *AluI* to *HpaI*, assembly of factor VIII genomic and cDNA clones from nucleotides -67 to 7,435 containing the entire open reading frame (nucleotides 2,277 to 3,795 are from a genomic clone, see Fig. 2); *HpaI* to *SstII*, pDLR1⁵⁰ fragment containing a portion of the Hepatitis B surface antigen gene 3' untranslated region; *SstII* to *SstI*, pFD11²⁵ fragment containing the remaining 3' untranslated region of the Hepatitis B surface antigen gene, a second SV40 early promoter and the 5' half of the dihydrofolate reductase gene (the *ClaI* site in this fragment has been removed by cutting, filling with Klenow polymerase, and religating); *SstI* to *SalI* pFD11²⁵ fragment containing the 3' half of the dihydrofolate reductase gene and pBR322 sequences (the *SstII* site in this fragment has been removed by cutting, trimming with Klenow polymerase and religating).

protein sequence and cloned DNA sequences, as well as the X-chromosome location for the cloned DNA, provided strong evidence that a portion of the factor VIII gene had been isolated. Characterization of the complete gene shows that this initially cloned region is located in the 3' half and corresponds to exon 16 (Ref. 16; Figs 2, 3).

cDNA cloning

The genomic factor VIII sequences thus identified were used to screen by blot hybridization nearly 80 different epithelial, connective and haematopoietic tissues and tumour-derived cell lines for a source of factor VIII mRNA. RNA from only one cell line examined in this fashion, the T-cell hybridoma line AL-7 (provided by D. Martin), exhibited a hybridizing mRNA species (Fig. 2a). The size of this mRNA, approximately nine kilobases (kb), is sufficient to code for a protein of $M_r \sim 300,000$, the reported size of factor VIII⁴. We have also detected factor VIII mRNA in samples of human liver; intact factor VIII RNA, however, was obtained more readily from cultured AL-7 cells. Factor VIII RNA from both sources share the same 5' cap site¹⁶.

A series of overlapping clones extending 5' from the exon 16 hybridizing region was isolated from specifically primed cDNA clones derived from AL-7 cell RNA (see Fig. 2 legend). We identified factor VIII cDNA clones (Fig. 2) in low abundance (about 1 in 100,000) by two routes: λ BRB22 cloning, using hybrid selection for the enrichment of sequences, and λ gt10 cloning (when hybrid enrichment could be dispensed with). The characterization of a 3,106 bp exon from genomic cloning¹⁶ allowed us to jump forward, using a cDNA primer based on the sequence of the 5' part of this exon. The DNA sequence of

the clone λ 13.2 indicated that the N-terminal coding region of factor VIII had been isolated.

The first 109 bases contain translational termination codons in all three reading frames, followed by an ATG initiation triplet then an open reading frame in the set of overlapping clones extending for 5,586 bp to the 3' end of exon 16 (Fig. 3). The deduced protein sequence reveals a 19 amino acid stretch beginning at the initiator methionine typical of a secreted protein leader sequence¹⁷. The salient features of the factor VIII presequence are two charged residues bordering a core of 10 hydrophobic amino acid residues (Fig. 3). The leader sequence is followed by a region virtually identical to the 25 amino acid N-terminal sequence determined for the M_r 210,000 and 95,000 fragments of purified factor VIII⁵.

cDNA clones containing the 3' end of the factor VIII gene and overlapping existing clones were isolated by oligo(dT) priming of RNA followed by specifically primed second strand synthesis and cloning in λ gt10 (Fig. 2). We obtained two clones containing the stop codon followed by an 1,805 bp 3' untranslated region with the polyadenylation signal AATAAA¹⁸ and different lengths of poly(A) tail at the same location. The complete set of overlapping clones was sequenced and ultimately reassembled to express factor VIII protein from recombinant DNA clones.

cDNA sequence

The complete factor VIII sequence (Fig. 3) predicts a polypeptide of 2,351 amino acids including a 19 amino acid signal peptide. The mature protein contains 2,332 amino acids of calculated $M_r \sim 265,000$. Considering potential glycosylation, this approximates to the reported M_r 330,000 native protein determined by SDS polyacrylamide gel electrophoresis⁴. The DNA derived amino acid sequence (Fig. 3), in combination with direct protein sequence analysis, not only confirms the presence of a very large single-chain form of the protein, but reveals the organization and identity of the plasma derived and thrombin cleaved fragments observed for the protein^{4,5,19-21}.

A computer-aided search for internal homologies in the factor VIII sequence reveals three types of domains. These consist of a triplicated region of 330-380 amino acids (A domains; positions 1-329, 330-711 and 1,649-2,019); a unique B domain of about 925 amino acids (712-1,648) and a duplicated region of about 160 amino acids (C domain; 2,020-2,172 and 2,173-2,332) (refs 5, 16). The regions are arranged A1-A2-B-A3-C1-C2. The A units have ~30% pairwise amino acid homology overall and the two C units 37%. Both the A and C repeats show considerable conservation of cysteine residues, indicating that structural similarity has been maintained within the repeated domains⁵. Most of the potential asparagine-linked glycosylation sites predicted by the sequence (19 of 25) are located in the B region.

The factor VIII protein sequence has remarkable homology with the sequence of ceruloplasmin, a copper containing plasma protein²². Ceruloplasmin consists of three successive repeats of 350 amino acids having a pairwise homology of ~40% (ref. 23) and shares ~30% homology with the triplicated A domains of factor VIII. Cysteine residues are found in the same relative location on the ceruloplasmin repeat unit as in the A regions of factor VIII, further affirming the relation of these proteins (see ref. 5, for further discussion).

Sequence expression

To demonstrate that the DNA sequences isolated encode factor VIII, the complete 7 kb protein-coding sequence was assembled from portions of the overlapping cDNA and genomic clones shown in Fig. 2, and inserted into a plasmid capable of transcribing heterologous sequences upon transfection into mammalian cell cultures. The structure of this expression plasmid, pSAT.8c1, is shown in Fig. 4. Factor VIII sequences were placed between a tandem SV40 early promoter/adenovirus-2 major late promoter and the hepatitis B virus surface antigen gene polyadenylation sequences²⁴. In addition, the plasmid contains a murine

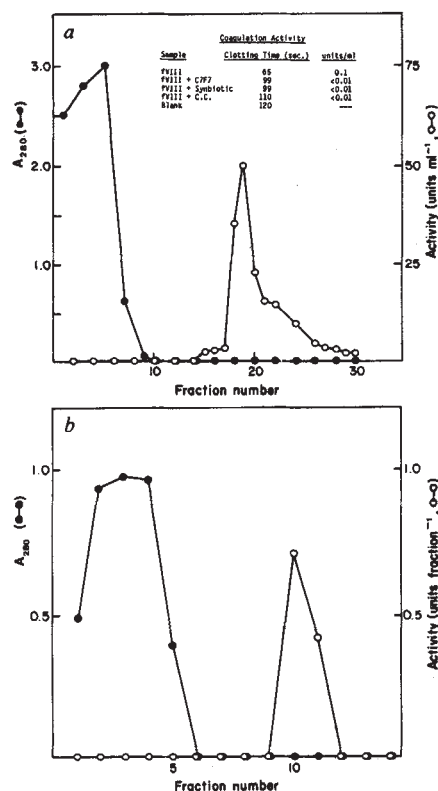


Fig. 5 Immunoaffinity chromatography of recombinant factor VIII. *a*, Chromatography on C7F7-Sepharose; *b*, chromatography on von Willebrand factor-Sepharose.

Methods: Serum free culture media of transformed cell line 62.2 containing 400 units of factor VIII was first adsorbed onto DEAE-sepharose (Pharmacia). The gel was washed subsequently with 0.05 M imidazole pH 6.9 containing 0.15 M NaCl, 0.01 M CaCl₂ and 0.02 M glycine ethyl ester; factor VIII was eluted with the same buffer containing 0.1 M CaCl₂. Material obtained from DEAE-sepharose was then passed through a factor VIII monoclonal antibody column. The column was prepared by covalently linking 2 mg factor VIII monoclonal antibody C7F7 to 1 ml of Affi-gel 10 (Bio-Rad). After loading, the column was washed sequentially with 20 volumes of 0.02 M imidazole pH 6.4 containing 0.15 M NaCl and 0.01 M CaCl₂ and 10 volumes of the same buffer containing 0.5 M NaCl. Factor VIII activity was eluted with 0.02 M imidazole pH 6.4 containing 1 M KI and 20% ethylene glycol and measured by the Coatest assay (Table 1). This figure depicts chromatography performed with the monoclonal antibody C7F7, known to react with the M_r 80,000 plasma-derived factor VIII protein fragment. Comparable results were obtained with a column using the C8 monoclonal antibody which binds to the M_r 210,000 and 90,000 fragments of factor VIII. The table inset of *a* shows the neutralization of coagulant activity by factor VIII antibodies. Partially purified recombinant factor VIII was dialysed into 0.05 M Tris pH 7.4 containing 0.15 M NaCl, then diluted with the same buffer containing 0.2% BSA. 50 µl aliquots of this sample were preincubated alone, with human factor VIII monoclonal antibodies (5 µg of either C7F7 or a human factor VIII monoclonal purchased from Synbiotics, Inc.), or with 5 µl of serum derived from a patient with a high factor VIII antibody titre⁴⁰ (C.C.) for 5 min at 24 °C. Factor VIII activity was then determined by coagulation assay. Briefly, 50 µl platelin (General Diagnostics) was incubated with 50 µl factor VIII deficient plasma for 8 min at 37 °C. 50 µl sample was then added and after 30 s at 37 °C, 50 µl 25 mM CaCl₂ was added and the clotting time measured. Supernatant from 62.2 cell cultures containing ~2 units of factor VIII activity was passed through a von Willebrand factor-sepharose column at 30 ml h⁻¹. The von Willebrand factor used here was purified as described previously³¹. von Willebrand factor-sepharose was prepared by covalently linking the protein to Affi-gel 10 (1 mg ml⁻¹ gel). After loading the column, it was washed with 5 volumes 0.05 M imidazole pH 6.9 containing 0.15 M NaCl, 0.01 M CaCl₂ and 0.02 M glycine ethyl ester. Factor VIII activity was eluted with the same buffer containing 0.25 M CaCl₂ and fractions assayed with the Coatest assay.

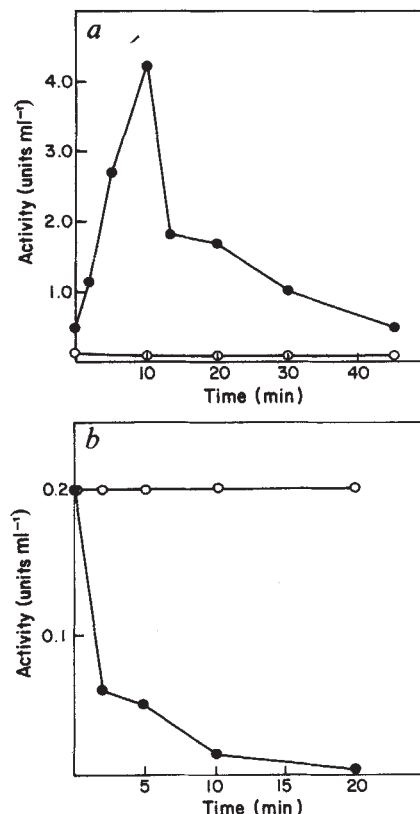


Fig. 6 Effect of thrombin and protein C on factor VIII activity. *a*, Recombinant factor VIII obtained from immunoaffinity chromatography was dialysed into 0.05 M Tris pH 7.4 buffer containing 0.15 M NaCl and diluted to a concentration of 0.1 units ml⁻¹ with the same buffer containing 0.2% BSA (TBS buffer). Aliquots of this sample were then incubated either with thrombin at 0.01 units ml⁻¹ (closed circles) or without thrombin (open circles) for the times shown at 37 °C. At the given time points, factor VIII activity was determined by the coagulation assay described in Fig. 5 legend. Thrombin at the concentration used here did not decrease background clotting times of samples in which factor VIII was omitted. *b*, Recombinant factor VIII was diluted into TBS buffer to a concentration of 0.2 units ml⁻¹. 5 µl rabbit brain cephalin (Sigma) and CaCl₂ to 5 mM were added to 50 µl aliquots of this sample. Subsequently these samples were incubated with activated protein C (provided by Walter Kiesel, University of Washington) at 0.4 mg ml⁻¹ (closed circles) or with no addition (open circles) for the times shown at 37 °C. Residual activity was then determined as described above.

dihydrofolate reductase cDNA transcribed by an SV40 early promoter to provide a selectable marker when expressed in mammalian cells²⁵. The plasmid was introduced into a hamster kidney cell line by the calcium-phosphate coprecipitation method^{25,26}; we chose one clone, 62.2, for further study.

To characterize the expression of factor VIII in the 62.2 cell line, RNA was extracted from cells and analysed by Northern blot hybridization. A predominant 9 kb RNA from 62.2 cells, but not from untransfected control cells, was found to hybridize both with a probe specific for factor VIII and with one containing hepatitis B surface antigen 3' sequences (data not shown). The intensity of the hybridizing signal with 62.2 RNA was 200 times greater than that observed with RNA prepared from the AL-7 cell line. Radioimmune assays for factor VIII were also performed on the 62.2 supernatant and cell lysate (Table 1). Monoclonal antibodies directed against the N-terminal M_r 210,000 fragment (C10) or the C-terminal M_r 80,000 fragment (C7F7) of serum-derived factor VIII detected a 300-fold increase in factor VIII cross-reacting activity in 62.2 cells over that found in control cells (Table 1).

Biochemical properties

The activation of natural factor X by factor IXa is dependent on factor VIII, which is the basis of the 'Coatest' assay we have used to estimate factor VIII in cell supernatants; we find significant activity in 62.2 supernatants but not in those from control cells not transfected (Table 1). Supernatant from 62.2 cells alone did not catalyse the cleavage of the chromogenic substrate. In addition, incubation of 62.2 cell supernatants with either of two monoclonal antibodies directed against the M_r 80,000 fragment of factor VIII (C7F7 and Synbiotic) completely neutralized the observed activity (Table 1).

To characterize further the factor VIII secreted by 62.2 cells, partially purified protein was tested for the ability to reduce the clotting time of plasma from patients with haemophilia A. Protein from 62.2 cell supernatants was affinity purified on a monoclonal antibody column under the same conditions used factor VIII isolated from human plasma (Fig. 5a). Diluted protein samples from the immunoaffinity column reduced the clotting time of plasma from a haemophilia A patient from 120 s to 65 s (equivalent to the clotting time for a 1:10 dilution of normal plasma). This activity was neutralized by the C7F7 and Synbiotic monoclonal antibodies directed against the M_r 80,000 fragment of factor VIII as well as by the patient-derived polyclonal antibody, C.C. (Fig. 5a).

Factor VIII circulates in plasma as a complex with the von Willebrand protein^{27,28}. Factor VIII/von Willebrand protein complexes are dissociated by high ionic strength or by calcium chloride containing buffers²⁹⁻³²; factor VIII derived from plasma can be bound to von Willebrand factor-Sepharose and eluted with CaCl_2 ³³. Recombinant factor VIII bound quantitatively to such an affinity column and could be completely eluted from the column with CaCl_2 , like plasma-derived factor VIII (Fig. 5b).

The coagulant activity of plasma-derived factor VIII is increased markedly by incubation with small amounts of the serine protease thrombin³⁴⁻³⁸. Thrombin activation probably results from the cleavage of specific peptide bonds in factor VIII⁵. After activation, continued incubation with thrombin leads to further degradation of the protein with a concomitant loss of coagulant activity. Recombinant factor VIII is affected by thrombin in a similar manner. The coagulation activity of the affinity purified protein from 62.2 cells was stimulated initially more than 40-fold by thrombin treatment (Fig. 6a). After longer incubation with thrombin, the activity substantially decreased (Fig. 6a). Plasma-derived factor VIII has also been shown to be inactivated by activated protein C^{34,39}, a vitamin K-dependent serine protease found in plasma, possibly functioning as a physiological regulator of coagulation. Incubation of the affinity purified 62.2 factor VIII with activated protein C resulted in the rapid loss of coagulation activity, as expected (Fig. 6b).

Conclusions

We have demonstrated by several criteria that the DNA sequences isolated here encode human factor VIII. Thus, genomic blots of male and 4X DNA hybridize with the 1:4 ratio expected for an X-linked gene. Moreover, the amino acid sequence derived from factor VIII preparations and the primary structure of the protein predicted by the cloned sequence are essentially identical⁵. Indeed, purified factor VIII preparations from plasma usually consist of a set of fragments between M_r 90,000 and 210,000 with very similar tryptic peptide maps and another, with M_r of 80,000, with a distinct map. Comparison of tryptic peptide sequences obtained from both classes of plasma-derived polypeptides⁵ with the DNA sequence of the cloned gene demonstrates that the M_r 210,000 and 80,000 polypeptides represent the N-terminal and C-terminal portions, respectively, of a single 2,332 amino acid mature polypeptide gene product. This protein presumably corresponds to the non-reducible M_r 330,000 species observed in purified factor VIII preparations⁴. The discrepancy between the molecular mass of single-chain plasma-derived factor VIII and that calculated for the cloned protein may be accounted for by the 25 potential

asparagine-linked glycosylation sites predicted by the latter. The peak of factor VIII activity found upon thrombin incubation correlates with the enhanced presence of the M_r 90,000 and a 70,000 fragment²¹. We have demonstrated that the 90,000 fragment corresponds to the amino terminal portion of the full-length molecule, that it derives from carboxy-terminal cleavage of the 210,000 fragment and that it contains the A1 and A2 repeated domains. The 70,000 species derives from the carboxy-terminus of the molecule following cleavage of the 80,000 fragment, including the A3, C1 and C2 domains⁵. Thus, removal of the unique B domain which connects the 90,000 and 80,000 fragments is associated with activation of the protein.

The recombinant factor VIII from the supernatants of cells transfected with the reconstructed gene corrects the coagulation defect of factor VIII-deficient plasma. A more stringent test of the recombinant protein was provided by the 'Coatest' assay, which measures the ability of factor VIII to accelerate the rate at which factor IXa proteolytically activates factor X in part of the intrinsic pathway of blood coagulation. The ability of several antibodies directed against factor VIII to neutralize these activities rules out the possibility that the cloned protein corresponds to other factors, such as the tissue factor/factor VII complex of the extrinsic coagulation pathway, also known to activate factor X.

It is anticipated that the availability of recombinant factor VIII from serum-free cell supernatants will further aid the study of structural modifications made on factor VIII by other components of the coagulation cascade. Such studies will be facilitated by the opportunity to create alterations in the structure, processing and activity of factor VIII by *in vitro* mutagenesis and heterologous expression of the cloned DNA sequences. Moreover, it is hoped that recombinant factor VIII will provide a safer and more plentiful therapy for individuals afflicted with haemophilia A.

We thank D. Hansen, B. Baer, U. Edman, T. Goralka, A. Kleiss, D. Leung, B. O'Connor and J. Souder for contributing to the work described here. Identification of cell line producing factor VIII RNA, cloning and assembly of factor VIII cDNA was performed by D. Capon. We also thank Genentech's Assay Development Group, Monoclonal Antibody Laboratory and the Organic Chemistry Laboratory. We appreciate the helpful discussions and continued support of D. Goeddel, A. Levinson, D. Martin, A. Ullrich, G. Hooper and R. Swanson and thank J. Arch and A. Gray for preparing the manuscript and figures. This work was supported by Genentech, Inc. and Speywood Laboratories Ltd.

Received 9 August; accepted 27 September 1984.

1. Davie, E. W. *et al.* *Rev. Biochem.* **44**, 799-829 (1975).
2. Jackson, C. M. & Nemerson, Y. A. *Rev. Biochem.* **49**, 767-811 (1980).
3. Fay, P. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 7200-7204 (1982).
4. Rotblat, F. *et al.* *Thromb. Haemostasis* **50**, 108 (1983).
5. Vehar, G. A. *et al.* *Nature* **312**, 337-342 (1984).
6. Zimmerman, T. S. & Meyer, D. in *Haemostasis and Thrombosis* (eds Bloom, A. L. & Duncan, P. T.) 111-123 (Churchill Livingstone, New York, 1981).
7. Meili, E. O. & Straub, P. W. *Thromb. Diath. haemorrh.* **24**, 161-174 (1970).
8. Stel, H. V., van der Kwast, T. H. & Veerman, E. C. I. *Nature* **303**, 530-532 (1983).
9. Kelly, D. A., Summerfield, J. A. & Tuddenham, E. G. D. *Br. J. Haemat.* **56**, 535-543 (1984).
10. Exner, T., Rickard, K. A. & Kronenberg, H. *Thromb. Res.* **32**, 427-436 (1983).
11. Anderson, S. & Kingston, I. B. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6838-6842 (1983).
12. Ullrich, A. *et al.* *EMBO J.* **3**, 361-364 (1984).
13. Jaye, M. *et al.* *Nucleic Acids Res.* **11**, 2325-2335 (1983).
14. Breathnach, R. & Chambon, P. A. *Rev. Biochem.* **50**, 349-383 (1981).
15. Sharp, P. A. *Cell* **23**, 643-646 (1981).
16. Gitschier, J. *et al.* *Nature* **312**, 326-330 (1984).
17. Perlman, D. & Halvorson, H. O., *J. molec. Biol.* **167**, 394-409 (1983).
18. Proudfoot, N. & Brownlee, G. *Nature* **252**, 359-362 (1981).
19. Fay, P. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 7200-7204 (1982).
20. Fulcher, C. A. & Zimmerman, T. S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1648-1652 (1982).
21. Fulcher, C. A., Roberts, J. R. & Zimmerman, T. S. *Blood* **61**, 807-811 (1983).
22. Takahashi, N., Ortel, T. & Putnam, F. *Proc. natn. Acad. Sci. U.S.A.* **81**, 390-394 (1984).
23. Takahashi, N. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 115-119 (1983).
24. Crowley, C. W., Liu, C. C. & Levinson, A. D. *Molec. Cell Biol.* **3**, 44-45 (1983).
25. Simonsen, C. C. & Levinson, A. D. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2495-2499 (1983).
26. Wigler, M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **76**, 1373-1376 (1979).
27. Schmer, G. *et al.* *J. biol. Chem.* **247**, 2512-2521 (1972).
28. Legaz, M. E. *et al.* *J. biol. Chem.* **248**, 3946-3955 (1973).
29. Tran, T. H. & Duckert, F. *Thromb. Haemostasis* **50**, 547-551 (1983).
30. Mikaelsson, M. E., Forsman, N. & Oswaldsson, U. M. *Blood* **62**, 1006-1015 (1983).
31. Weiss, H. & Hoyer, L. *Science* **182**, 1149-1151 (1973).
32. Baugh, R. *et al.* *Biochim. biophys. Acta* **371**, 360 (1974).
33. Fass, D. N., Knutson, G. J. & Katzmann, J. A. *Blood* **59**, 594-600 (1982).
34. Vehar, G. A. & Davie, E. W. *Biochemistry* **19**, 401-410 (1980).

35. Rapaport, S. I., Schiffman, S., Patch, M. J. & Ames, S. B. *Blood* **21**, 221-236 (1963).
36. Macfarlane, R. G., Biggs, R., Ash, B. J. & Denson, K. W. *Br. J. Haemat.* **10**, 530-538 (1964).
37. Chavin, S. I. *Am. J. Hemat.* **16**, 297-306 (1984).
38. Fulcher, C. A. *et al. Blood* **63**, 486-489 (1984).
39. Marlar, R. A., Kleiss, A. J. & Griffin, J. H. *Blood* **59**, 1067-1072 (1982).
40. Rotblat, F. *et al. Thromb. Res.* **21**, 431-436 (1981).
41. Rotblat, F. *et al. J. Lab. clin. Med.* **101**, 736-746 (1983).
42. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
43. Kafatos, F. C., Jones, C. W. & Efstratiadis, A. *Nucleic Acids Res.* **7**, 1541-1552 (1979).
44. Capon, D. C. *et al. Nature* **304**, 507-513 (1983).
45. Hyunh, T., Young, R. & Davis, R. in *Practical Approaches in Biochemistry* (ed. Grover, D.) (IRL Press, Oxford, in the press).
46. Vieira, J. & Messing, J. *Gene* **19**, 259-268 (1982).
47. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
48. Liu, C. C., Simonsen, C. C. & Levinson, A. D. *Nature* **309**, 82-85 (1984).
49. Gingeras, R. *et al. J. biol. Chem.* **257**, 13475-13496 (1982).
50. Simonsen, C. C. & Levinson, A. D. *Molec. Cell Biol.* **3**, 2250-2258 (1983).
51. Weinstein, M. & Daykin, D. *Blood* **53**, 1095-1105 (1979).

Structure of human factor VIII

Gordon A. Vehar*, Bruce Keyt*, Dan Eaton*, Henry Rodriguez*,
Donogh P. O'Brien†, Frances Rotblat†, Herman Oppermann†, Rodney Keck*,
William I. Wood‡, Richard N. Harkins*, Edward G. D. Tuddenham‡,
Richard M. Lawn‡ & Daniel J. Capon*

* Department of Protein Biochemistry and † Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

‡ Haemophilia Centre, Academic Department of Haematology, Royal Free Hospital School of Medicine, London WC1N 1BP, UK

The deduced amino acid sequence of human factor VIII, obtained from the DNA sequence, predicts a mature polypeptide of 2,332 amino acids containing a triplicated domain structure. The polypeptide has 35% sequence homology with the copper-binding plasma protein, ceruloplasmin. Determination of the thrombin cleavage sites in plasma-derived factor VIII polypeptides allows prediction of the domains involved in the associated activation and inactivation of the protein.

PREPARATIONS of factor VIII/von Willebrand factor complex¹⁻¹² contain four closely related properties or activities: factor VIII coagulant activity, an antigen associated with the factor VIII coagulant activity, platelet adhesion promoting activity and an antigen precipitated by antisera raised against the purified complex (factor VIII-related protein). Factor VIII separated from the complex has associated trace amounts of protein, is unstable and consists of multiple polypeptide chains¹³⁻¹⁷, hindering detailed characterization studies.

The purification of human factor VIII by affinity to a monoclonal antibody directed against the coagulant activity of factor VIII¹⁸, has allowed characterization of the protein fragments of factor VIII or thrombin-activated factor VIII by partial amino acid sequence analysis. This sequence information has been used to isolate cDNA and genomic clones encoding human factor VIII^{19,20}. The protein sequence deduced from these clones, together with an analysis of the cleavage products associated with the activation of factor VIII by thrombin, allows the assignment of thrombin cleavage sites and the identification of most of the polypeptide fragments present in highly purified factor VIII preparations. The factor VIII sequence exhibits striking homology with the plasma copper-binding protein ceruloplasmin, suggesting novel biochemical activities for factor VIII as well as a role for metal ions other than calcium in the blood coagulation cascade.

Analysis of plasma factor VIII

Preparations of human factor VIII¹⁸ purified over 300,000-fold from plasma contained several proteins of relative molecular mass (M_r) 210,000-80,000 (Fig. 1A). These protein bands were not connected by disulphide links because samples analysed under non-reducing conditions gave a similar pattern (data not shown). To determine the relationship of these multiple polypeptide chains, we analysed them by tryptic peptide mapping. The preparation used in Fig. 1B contained a protein of M_r 240,000, producing a peptide map which did not show identity with the other proteins of the mixture (Fig. 1B, a), and has been found to be a von Willebrand factor subunit (data not shown). Peptides of M_r 90,000-210,000 all had a common tryptic map, indicating that they are derived from the same or closely related polypeptide chains (Fig. 1B, b-f). Furthermore, Western blot analysis of the factor VIII preparations demonstrated that a factor VIII-specific monoclonal antibody¹⁸ reacted with the M_r 90,000-

210,000 polypeptides (data not shown). Two very similar patterns generated by the proteins of M_r 80,000 and 70,000 had a different peptide map (Fig. 1B, g, h). These results demonstrate that the fragments of M_r 90,000-210,000 are structurally related and could be pooled and treated as one polypeptide chain. The protein represented by the band of M_r 80,000 (and 70,000 when present) was analysed separately.

The purified factor VIII preparations were fractionated by gel filtration on a TSK 4000SW HPLC column; analysis of the resulting fractions demonstrated the effective separation of M_r 240,000 protein, the polypeptides of M_r 90,000-210,000 and the M_r 80,000 fragment (Fig. 2B). We performed amino acid sequence analysis on peptides generated from the M_r 80,000 protein, the polypeptide pool of M_r 90,000-210,000 and a fragment of M_r 90,000 from limited thrombin digestion of the M_r 90,000-210,000 pool. After digestion of each sample with trypsin, the resulting peptides were separated by reverse-phase HPLC and sequenced. The peptide sequence AWAYFSDVDLEK, used to prepare synthetic DNA probes identifying factor VIII genomic DNA clones, is indicated in Fig. 2C.

Structure of factor VIII protein

The molecular cloning of the entire factor VIII coding region is described in an accompanying paper¹⁹. The 2,351-amino acid sequence for factor VIII, deduced from the nucleotide sequence of these clones, is shown in Fig. 3. The first 19 amino acids of the sequence comprise the signal sequence for factor VIII, based on peptide sequence analysis of a fragment derived from the M_r 90,000-210,000 polypeptide pool. The N-terminal sequence of this M_r 30,000 fragment, obtained as a thrombin digest product of the M_r 90,000-210,000 pool, is identical to the first 12 amino acids which follow the predicted factor VIII leader sequence (see Fig. 5). This presequence exhibits a core of 10 hydrophobic amino acids flanked by two charged residues, a structure which conforms to that observed for the leader sequences found in most secreted proteins²¹. The mature protein contains 2,332 amino acids (calculated M_r 264,763).

The availability of the complete factor VIII sequence reveals the organization and identity of the tryptic peptides obtained from the pools of separated plasma-derived factor VIII fragments. Essentially all tryptic peptide sequences determined from the M_r 90,000-210,000 protein pools are located in the amino-

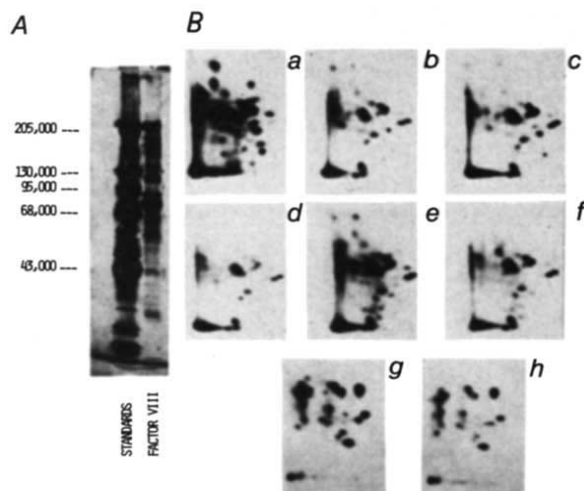


Fig. 1 A, Purified human factor VIII analysed by SDS-polyacrylamide gel electrophoresis. B, Two-dimensional tryptic mapping of factor VIII polypeptide chains. The resulting tryptic patterns of proteins of M_r : a, 240,000; b, 210,000; c, 170,000; d, 150,000; e, 120,000; f, 100,000; g, 80,000; and h, 70,000 are shown.

Methods: The purified protein¹⁸ was reduced and analysed in a 5–10% polyacrylamide gradient gel in the presence of SDS by the procedure of Laemmli³⁷. The molecular weights of the protein standards are shown (myosin, β -galactosidase, phosphorylase B, bovine serum albumin and ovalbumin). The bands were detected using the silver stain procedure of Morrissey³⁸. 1 μ g of a factor VIII preparation was denatured in 1% SDS and labelled with 300 μ Ci 125 I for 15 min using the iodobead procedure³⁸. Labelled polypeptides were located on dried SDS-polyacrylamide gels by autoradiography and digested by incubation of gel slices with 10 μ g trypsin in 0.1 M ammonium bicarbonate buffer for 6 h at 37°C. After repeated lyophilizations, peptides were dissolved in 8.8% formic acid and a portion was subjected to thin-layer electrophoresis in the same buffer (400 V for 45 min) on pre-coated TLC-cellulose plates (E. Merck, Darmstadt, FRG). For the second dimension, peptides were separated by ascending chromatography in *n*-butanol/pyridine/glacial acetic acid/water, 75:50:15:60 (v/v). The plates were then subjected to autoradiography.

terminal half of the molecule, whereas those sequences obtained for the M_r 80,000 fragment are found at the carboxy-terminus of the factor VIII sequence (unpublished results). The most carboxy-terminal tryptic peptides identified for the M_r 90,000–210,000 pool gave the sequences GEFT and -QEE, beginning at positions 1,155 and 1,194, respectively. This shows that the M_r 210,000 fragment consists of a protein of $M_r \geq 135,000$ containing 14 potential asparagine-linked glycosylation sites. The location of the M_r 80,000 fragment of factor VIII is delineated by two peptide sequences which define a stretch of ~680 amino acids. The first of these was obtained from the amino-terminal sequence of the M_r 80,000 fragment beginning at position 1,649 (see Fig. 5), whereas the second corresponds to a tryptic peptide, MEVLGCEAQDL, 12 amino acids from the C-terminus predicted by the DNA sequence. Thus, there is no significant removal of C-terminal sequences from the plasma-derived molecule. The failure to recover tryptic peptide sequences from the region between position 1,200 and the M_r 80,000 fragment is probably due to the relatively low concentration of the M_r 210,000 species in the M_r 90,000–210,000 fragment pool. This position should therefore be considered the minimal C-terminal extent of the M_r 210,000 protein.

Computer-aided analysis of the factor VIII protein sequence revealed two types of internal homology: the first consists of a triplicated segment (A domain) found at positions 1–329, 380–711 and 1,649–2,019 of the mature polypeptide (Fig. 4); the second and third domains of the triplication are separated by a region of 983 amino acids (B domain) extremely rich in potential asparagine-linked glycosylation sites. In addition, an unrelated duplication of 150 amino acids is found at the C-terminus of the molecule (C domain). The A domains have ~30% amino acid homology, whereas the C domains are ~40% homologous. Most

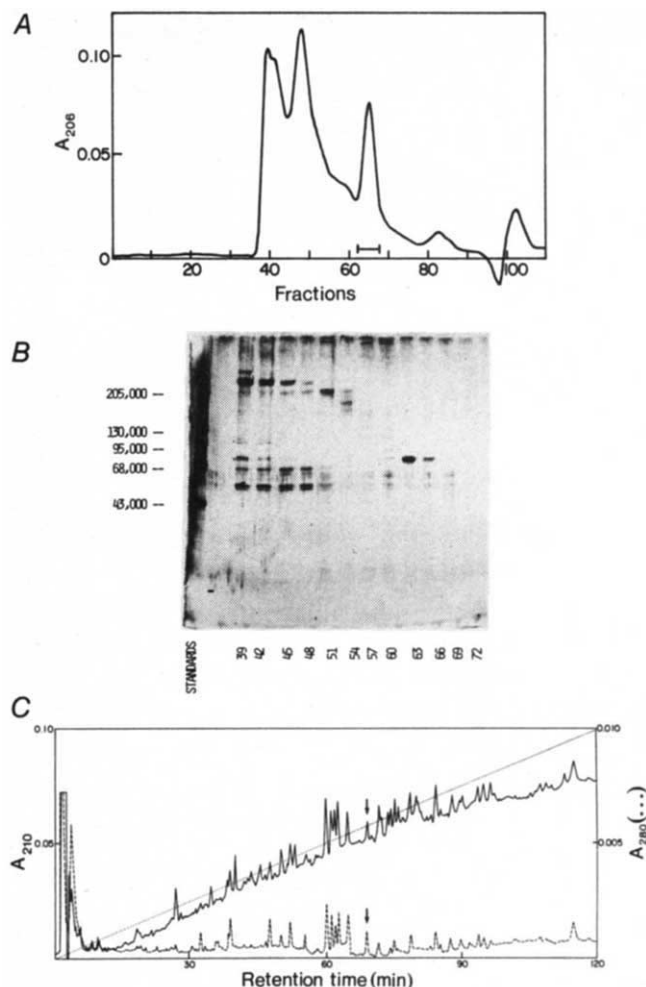


Fig. 2 Purification of factor VIII polypeptides. A, Fractionation of proteins by TSK 4000 HPLC. B, SDS-polyacrylamide gel analysis of fractions from TSK 4000 chromatography. C, Reverse-phase HPLC separation of tryptic peptides of M_r 80,000 protein.

Methods: A, Human factor VIII preparations¹⁸ were dialysed into 1% ammonium bicarbonate containing 0.1% SDS. The samples were lyophilized and stored at -20°C until use. The samples were reconstituted in distilled water and applied to a TSK 4000 SW column (0.75 $\times$ 50 cm; Alltech Associates, Deerfield, Illinois) equilibrated with 0.1 M sodium phosphate buffer, pH 7.0, containing 0.1% SDS. Samples of ~0.15–0.25 ml were injected and the column developed isocratically at a flow rate of 0.5 ml min⁻¹. Absorbance was monitored at either 206 nm or 280 nm and fractions (0.2 ml) collected. Fractions across the profile were reduced and analysed by SDS-polyacrylamide gel electrophoresis as described in Fig. 1 legend. The protein of M_r 80,000 was pooled for tryptic digestion as shown in A. C, The TSK 4000 SW purified protein of M_r 80,000 (0.8 nM) was dialysed under a nitrogen atmosphere overnight against 0.36 M Tris-HCl, pH 8.6, containing 8 M urea, 3.3 mM EDTA and 10 mM dithiothreitol, DTT (final vol. 1.5 ml). The protein was alkylated by adding 15 μ l of 5 M iodoacetic acid (dissolved in 1 M NaOH). The reaction was allowed to proceed for 35 min at room temperature in the dark and was quenched by adding DTT to a final concentration of 100 mM. The protein solution was dialysed against 8 M urea in 0.1 M ammonium bicarbonate for 4 h. The urea dialysis solution was changed over a period of 24 h to gradually reduce the urea concentration to a final level of 0.5 M. Trypsin was added at a weight ratio of 1:50 at 37°C for 12 h. HPLC separation of the resulting tryptic peptides was performed on a high-resolution Synchropak RP-P C-18 column (0.46 $\times$ 25 cm; 10 μ). The column was developed with a gradient of acetonitrile (1–70% in 200 min) in 0.1% trifluoroacetic acid. Absorbance was monitored at 210 nm and 280 nm. Each peak was collected and stored at 4°C until subjected to sequence analysis in a Beckman spinning cup sequencer with on-line phenylthiohydantoin amino acid identification⁴⁰. The arrow identifies the peptide (AWAYFSDVDLEK) resulting in identification of a factor VIII genomic clone^{19,20}.

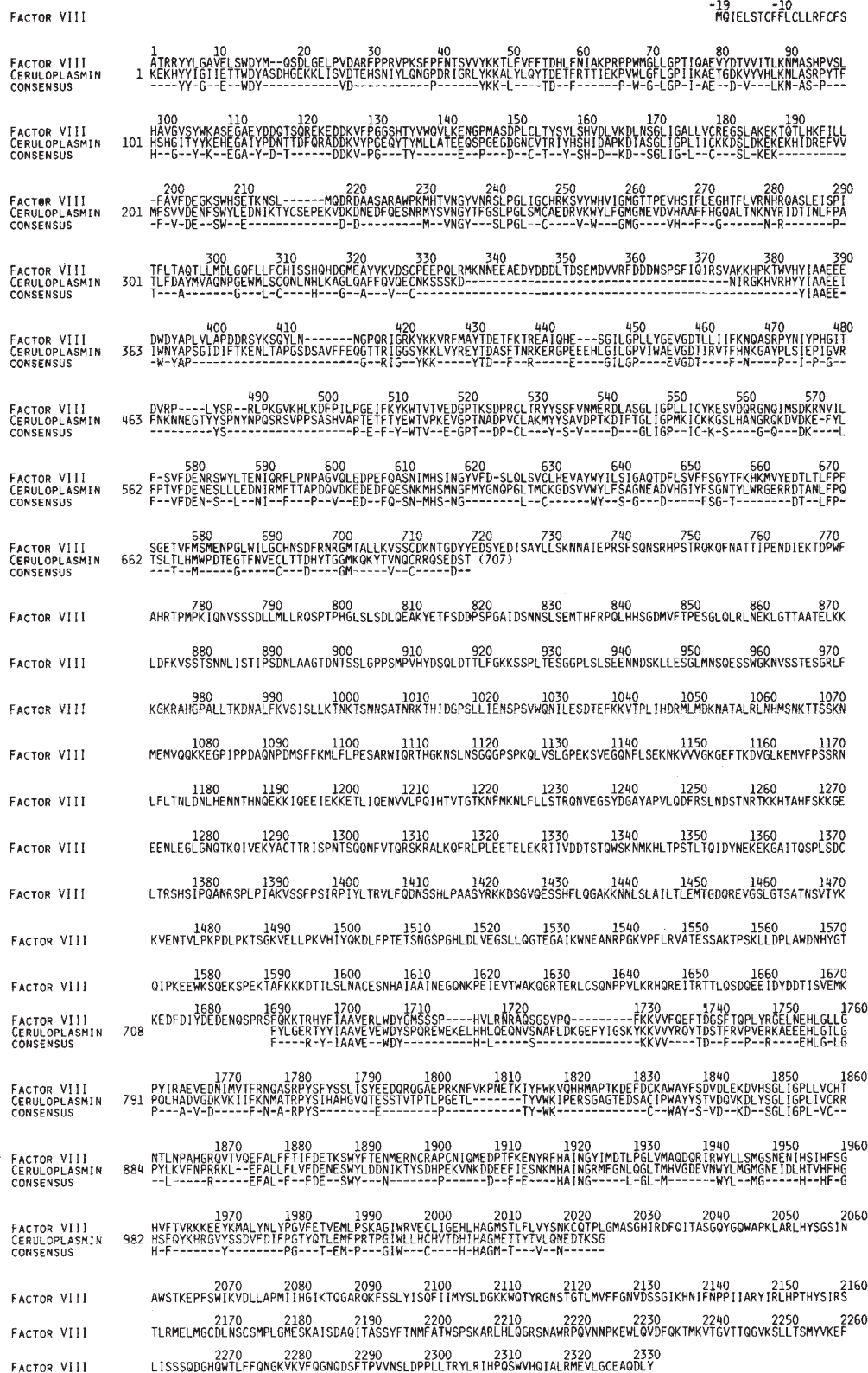


Fig. 3 Homology of human factor VIII with human ceruloplasmin. The deduced factor VIII amino acid sequence derived from cDNA and genomic clones^{19,20} is compared with that of human ceruloplasmin²². The consensus line shows residues which are identical in the two proteins. The numbering above the sequences is that of factor VIII. The number preceding the ceruloplasmin line represents the numbering of the amino acid sequence beginning the line. The single-letter notation for amino acids is used: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.

of the 23 cysteine residues of the mature polypeptide are clustered in the A and C domains and occupy similar positions (Fig. 6), suggesting that the structures of both repeated domains of factor VIII reflect conserved disulphide bonding arrangements.

The A domains of the factor VIII protein show striking homology with the copper-binding plasma protein ceruloplasmin (Fig. 3). Amino acid sequence analysis of ceruloplasmin has revealed a structure consisting of three contiguous domains sharing ~30% homology²²⁻²⁴. The triplicated domains of factor VIII and ceruloplasmin exhibit a pairwise homology of 30%

(Fig. 4). Although the B domain has no substantial homology with any known sequence, the C domain shares 20% amino acid homology with the discoidin lectins from *Dictyostelium*²⁵.

Ceruloplasmin contains six copper atoms in three distinct types of coordination: two of type 1, one of type 2 and three electron paramagnetic resonance-nondetectable type 3 copper ions²⁶. The type 1 copper ions are thought to bind to the carboxy-terminal portion of the domains of ceruloplasmin (domain residues 240–350; Fig. 4) based on sequence homology with the type 1 copper-binding protein plastocyanin²⁷. The four amino acid side chains proposed as the ligands for the type 1

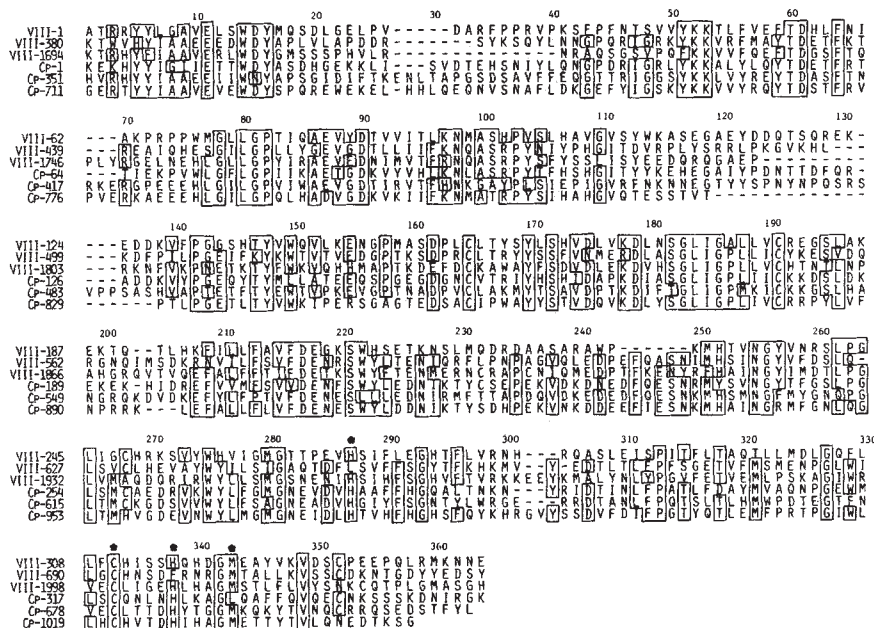


Fig. 4 Domain homology of human factor VIII and human ceruloplasmin²². Identical residues in four of the six domains are boxed. The protein is indicated to the left of the lines (VIII, factor VIII; Cp, ceruloplasmin) with the associated number indicating the position in the sequence of the first amino acid of the line. A domain residue numbering system is shown above the sequences. * Residues believed to be the ligands for type I copper (based on homology to plastocyanin²⁷).

copper atoms of ceruloplasmin are found in the first and third A domains of factor VIII (histidines 287 and 338, cysteine 333 and methionine 343; see Fig. 4). The conservation of the copper-ligand residues found in ceruloplasmin strongly suggests similar metal binding characteristics for factor VIII.

Thrombin cleavage

The coagulant activity of factor VIII is increased markedly by treatment with catalytic amounts of the serine protease thrombin²⁸. Thrombin activation of factor VIII is associated with a series of polypeptide cleavages¹³⁻¹⁷. Further incubation with thrombin leads to degradation of the protein with a concomitant loss of coagulant activity¹³⁻¹⁷. To understand the structural basis for these observations, the separated pools of factor VIII fragments were subjected to thrombin digestion and the resulting products characterized by SDS-polyacrylamide gel electrophoresis. Thrombin digestion of the M_r 80,000 protein resulted in a product of M_r 73,000 (Fig. 5A). Treatment of the polypeptide pool of M_r 90,000-210,000 led initially to the appearance of two bands of M_r 43,000 and 50,000 (Fig. 5A). Longer incubation with thrombin resulted in the conversion of the M_r 50,000 fragment to polypeptides of M_r 30,000 and 20,000 (data not shown). The M_r values of these thrombin digest fragments are similar to those generated by thrombin treatment of native factor VIII preparations^{13,15,16}. Amino-terminal sequence analysis was performed on the separated protein chains before and after thrombin digestion. The resulting sequences are compared with the corresponding amino acid sequences deduced from the factor VIII cDNA sequence¹⁹ in Fig. 5B; also shown are the potential cleavage sites found at the amino-terminus of the M_r 80,000 protein and that which separates the M_r 90,000 fragment from the carboxy-terminal portion of the M_r 210,000 protein. The sequence surrounding the latter potential cleavage site (position 740) is similar to the amino-terminal sequence of the M_r 70,000 protein (-PRSF---RH-) (Fig. 5B). There is no other consensus sequence that would predict the specificity of thrombin cleavage. A homologous stretch preceding the thrombin cleavage sites for the M_r 43,000 and 73,000 proteins is observed (Fig. 5B), but whether this homology determines thrombin specificity or simply reflects the internal duplication is uncertain. The most consistent sequence found at thrombin cleavage sites within factor VIII is an arginine residue followed by either serine or alanine. Other such sequences (-RS- or -RA-) do occur within the protein but are not cleaved, suggesting the possible involvement of secondary structure in thrombin specificity. The cleavage that frees the M_r 80,000 protein occurs at an arginine-glutamic acid sequence, probably not a thrombin-generated cleavage site. This cleavage occurs

quickly and is complete within the time required for isolation of the protein. The precursor factor VIII protein therefore may be cleaved to free the M_r 80,000 polypeptide by a protease other than thrombin.

Discussion

The structure of factor VIII revealed by the amino acid sequence predicted from the cloned cDNA and the structural characterization of polypeptide fragments described here are summarized in Fig. 6. The size of the factor VIII precursor moiety is consistent with the reported isolation of single-chain M_r 330,000 protein from plasma¹⁸ and supports the notion that the protein circulates as a high-molecular weight form that is readily cleaved in plasma and/or during isolation to a series of degradation products.

The primary structure of factor VIII exhibits three distinct types of structural domain, including a triplicated region of ~330 amino acids (A domains), a unique region of 980 amino acids (the B domain) and a carboxy-terminal duplicated region of 150 amino acids (C domains), which are arranged in the order A1-A2-B-A3-C1-C2 (Fig. 6). The A domains of factor VIII show significant homology to ceruloplasmin, consisting also of a triplicated structure of three A domains but lacking both B and C domains²²⁻²⁴. Particularly striking is the clustering of cysteine residues at similar locations within related structural domains of factor VIII (Fig. 6). The determination of disulphide pairings for ceruloplasmin^{23,29} predicts two types of internal disulphide bonding arrangements for the A domains of factor VIII. The disulphide structure proposed for the C domains of factor VIII is based on the proposition that the disulphide linkages form between the two cysteine residues found in each domain. The large B domain which separates the second and third A domains of factor VIII contains only four cysteine residues, but the presence of 19 asparagine-linked glycosylation sites suggests that this region is extensively modified by carbohydrate addition.

The activation of factor X by factor IX_a in conjunction with factor VIII is known to require calcium ions. Factor IX_a and factor X both contain γ -carboxyglutamic acid residues which are thought to be involved in calcium binding. The protein-bound calcium ions mediate the interaction of these proteins with the phospholipid surface. The homology of factor VIII with ceruloplasmin suggests the possible involvement of copper or other metal ions in the role of factor VIII in factor X activation. One possibility for the role of such metal ion involvement is suggested by the binding of lanthanide ions by γ -carboxyglutamic acid residues³⁰. It is interesting to speculate that the potential copper-binding ligands of factor VIII interact with a metal ion jointly bound by the γ -carboxyglutamate

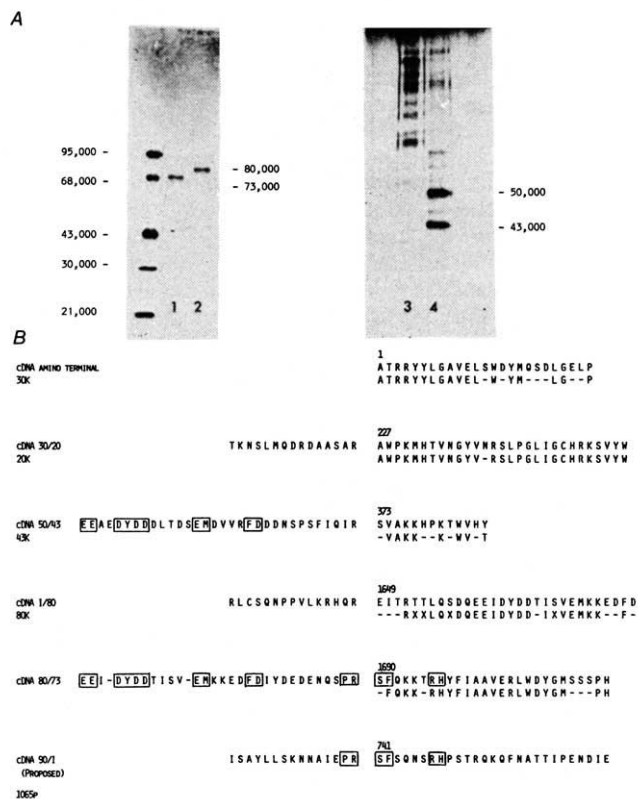
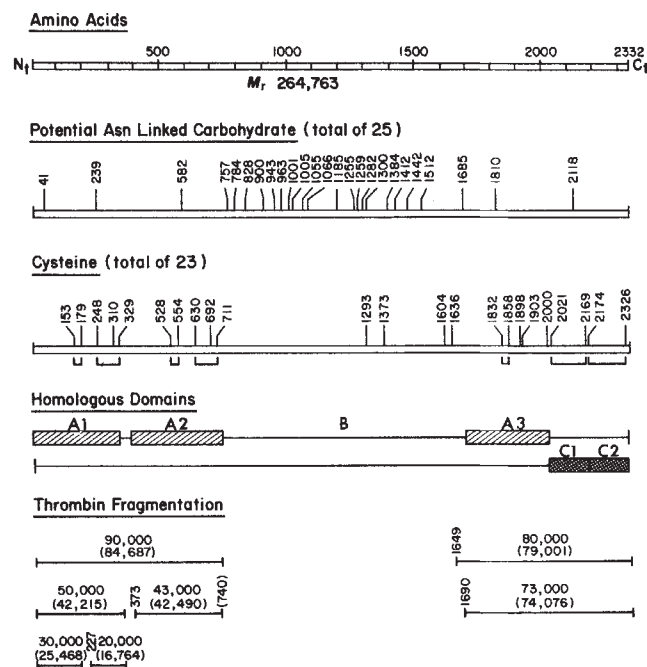


Fig. 5 **A**, Thrombin cleavage of separated factor VIII polypeptide fractions. **B**, Alignment of thrombin-generated amino termini with the deduced amino acid sequence¹⁹.

Methods: **A**, The factor VIII polypeptides were separated as described in legend to Fig. 2A. SDS was removed by dialysis of the fractions against 8 M urea solutions and urea removed by dialysis against 0.01 M Tris, pH 8.0. Thrombin was then added at weight ratios to a maximum of 1:20. The digestions were allowed to proceed at room temperature and the extent of cleavage was monitored by SDS-gel electrophoresis. Lanes 1 and 2 are the M_r 80,000 protein with and without thrombin, respectively; lanes 3 and 4 are the M_r 90,000–210,000 protein without and with thrombin, respectively. **B**, Thrombin digestion products were separated by preparative SDS-gel electrophoresis, electroeluted by the procedure of Hunkapiller *et al.*⁴¹ and analysed on a Beckman spinning cup sequencer⁴⁰. The M_r 80,000 protein was obtained by TSK separation as described for Fig. 2A. The thrombin cleavage site is indicated by a space in the cDNA deduced sequence. The number above the subsequent amino acid corresponds to the position of that residue in the linear sequence. The amino-terminal protein sequences for the various polypeptide chains are aligned under the translated gene sequence. —, Positions where no residue could be identified; X, positions where the wrong amino acid was determined. The relative molecular masses of the proteins separated are listed (30K is the amino terminus obtained for the gel-eluted polypeptide with a M_r of 30,000); solidi indicate cleavage products (for example, 50/43 indicates the cleavage which separates the M_r 50,000 and 43,000 species). Regions which share sequence homology are boxed.

residues of factors IX or X. In addition to copper transport and haemostatis, several enzymatic functions have been ascribed to ceruloplasmin, including ferroxidase activity, amino oxidase activity and superoxide dismutase activity^{31–33}. It will be important to determine whether any of these activities are associated with factor VIII.

The amino-terminal sequence of factor VIII thrombin fragments and the homology of factor VIII with ceruloplasmin provide insight into a functional purpose for the cleavages. Factor VIII isolated from plasma is usually degraded. In certain preparations, small amounts of a M_r 330,000 protein were observed when analysed on SDS-polyacrylamide gels run under non-reducing conditions; this protein was not observed when the factor VIII samples were reduced before electrophoresis. This reducible 330,000 may be a disulphide-linked, limited pro-



termini) from a single-chain circulating form. By analogy with factor V, the M_r 90,000 and 80,000 proteins would correspond to fragments D and E of factor V, respectively³⁴. These two fragments of factor V can be separated from the activation peptides and isolated as a functional two-subunit protein^{35,36}. Both subunits are required for factor V activity³⁵ and both may be required for factor VIII activity. A highly glycosylated intermediate region is cleaved from both proteins. Therefore, both factors V and VIII seem to be highly similar in structure, thrombin cleavage pattern and, presumably, function.

The studies described here provide a structural basis for defining the role of the diverse molecular forms of factor VIII in their interaction with other proteins of the coagulation cascade. The availability of complete factor VIII cDNA clones

capable of programming recombinant factor VIII synthesis in mammalian cell cultures¹⁹ will offer a unique opportunity to perform similar studies with the single-chain precursor molecule. The questions raised here concerning the relationship of processing events, structural domains and homology to ceruloplasmin with the biological function of factor VIII may be answered by studying structural changes introduced into the protein by modification of these cloned DNA sequences.

We acknowledge the support of Speywood Laboratories Ltd. We thank Dr John Bell, Alice Kleiss and William Henzel for laboratory assistance; Gary Hooper for his assistance in the coordination of the project between various groups and companies; and Robert A. Swanson and Dr David W. Martin Jr for their encouragement.

Received 13 August; accepted 27 September 1984.

1. Jackson, C. M. & Nemerson, Y. A. *Rev. Biochem.* **49**, 765-811 (1980).
2. Lundblad, R. L. & Davie, E. W. *Biochemistry* **3**, 1720-1725 (1964).
3. Hougie, C., Denson, K. W. E. & Biggs, R. *Thromb. Diath. haemorrh.* **18**, 211-222 (1967).
4. Hemker, H. C. & Kahn, M. J. P. *Nature* **215**, 1201-1202 (1967).
5. Barton, P. G. *Nature* **215**, 1508-1509 (1967).
6. Osterud, B. & Rapaport, S. I. *Biochemistry* **9**, 1854-1861 (1970).
7. Chuang, T. F., Sargeant, R. B. & Hougie, C. *Biochim. biophys. Acta* **273**, 287-291 (1972).
8. Ratnoff, O. D., Kass, L. & Lang, P. D. *J. clin. Invest.* **48**, 957-962 (1969).
9. Herschgold, E. J., Silverman, L., Davison, A. & Janszen, M. E. *J. Lab. clin. Med.* **77**, 185-205 (1971).
10. Schmer, G., Kirby, E. P., Teller, D. C. & Davie, E. W. *J. biol. Chem.* **247**, 2512-2521 (1972).
11. Legaz, M. E., Schmer, G., Counts, R. B. & Davie, E. W. *J. biol. Chem.* **248**, 3946-3955 (1973).
12. Shapiro, G. A., Anderson, J. C., Pizzo, S. V. & McKee, P. A. *J. clin. Invest.* **52**, 2198-2210 (1973).
13. Vehar, G. A. & Davie, E. W. *Biochemistry* **19**, 401-410 (1980).
14. Fass, D. N., Knutson, G. J. & Katzman, J. A. *Blood* **59**, 594-600 (1982).
15. Fulcher, C. A. & Zimmerman, T. S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1648-1652 (1982).
16. Knutson, G. J. & Fass, D. N. *Blood* **59**, 615-624 (1982).
17. Fulcher, C. A., Roberts, J. R. & Zimmerman, T. S. *Blood* **61**, 807-811 (1983).
18. Rotblat, F., O'Brien, D. P., Middleton, S. M. & Tuddenham, E. G. D. *Thromb. Haemostasis* **50**, 108 (1983).
19. Wood, W. I. *et al. Nature* **312**, 330-337 (1984).
20. Gitschier, J. *et al. Nature* **312**, 326-330 (1984).
21. Perlman, D. & Halvorson, H. O. *J. molec. Biol.* **167**, 394-409 (1983).

22. Takahashi, N., Ortel, T. L. & Putnam, F. W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 390-394 (1984).
23. Takahashi, N. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 115-119 (1983).
24. Dwulet, F. E. & Putnam, F. W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2805-2809 (1981).
25. Poole *et al. J. molec. Biol.* **153**, 273-289 (1981).
26. Ryden, L. & Bjork, I. *Biochemistry* **15**, 3411-3417 (1976).
27. Ryden, L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6767-6771 (1982).
28. Rapaport, S. I., Schiffman, S., Patch, M. J. & Ames, S. B. *Blood* **21**, 221-236 (1963).
29. Ryden, L. & Norder, H. *J. Chromatogr.* **215**, 341-350 (1981).
30. Sperling, R., Furie, B. C., Blumenstein, M., Keyt, B. & Furie, B. *J. biol. Chem.* **253**, 3898-3906 (1978).
31. Poulik, M. D. & Weiss, M. L. in *The Plasma Proteins* Vol. 2 (ed. Putnam, F. W.) 51-108 (Academic, New York, 1975).
32. Frieden, E. *Ciba Fdn Symp.* **79**, 93-124 (1980).
33. Goldstein, I. M., Kaplan, H. B., Edelson, H. S. & Weissman, G. *J. biol. Chem.* **254**, 4040-4045 (1979).
34. Nesheim, M. E., Foster, W. B., Hewick, R. & Mann, K. G. *J. biol. Chem.* **259**, 3187-3196 (1984).
35. Esmon, C. T. *J. biol. Chem.* **254**, 964-973 (1979).
36. Suzuki, K., Dahlback, B. & Stenflo, J. *J. biol. Chem.* **257**, 6556-6564 (1982).
37. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
38. Morrissey, J. H. *Analyt. Biochem.* **117**, 307-310 (1981).
39. Markwell, M. A. K. *Analyt. Biochem.* **125**, 427-432 (1982).
40. Rodriguez, H., Kohr, W. J. & Harkins, R. N. *Analyt. Biochem.* **140**, 538-547 (1984).
41. Hunkapiller, M. W., Lujan, E., Ostrander, F. & Hood, L. E. *Meth. Enzym.* **91**, 227-247 (1983).

Molecular cloning of a cDNA encoding human antihaemophilic factor

John J. Toole, John L. Knopf, John M. Wozney, Lisa A. Sultzman, Janet L. Buecker, Debra D. Pittman, Randal J. Kaufman, Eugene Brown, Charles Shoemaker, Elizabeth C. Orr, Godfrey W. Amphlett, W. Barry Foster, Mary Lou Coe, Gaylord J. Knutson*, David N. Fass* & Rodney M. Hewick

Genetics Institute, Inc., 225 Longwood Avenue, Boston, Massachusetts 02115, USA

* Hematology Research Section, Mayo Clinic/Foundation, Rochester, Minnesota 55905, USA

A complete copy of the mRNA sequences encoding human coagulation factor VIII:C has been cloned and expressed. The DNA sequence predicts a single chain precursor of 2,351 amino acids with a relative molecular mass (M_r) 267,039. The protein has an obvious domain structure, contains sequence repeats and is structurally related to factor V and ceruloplasmin.

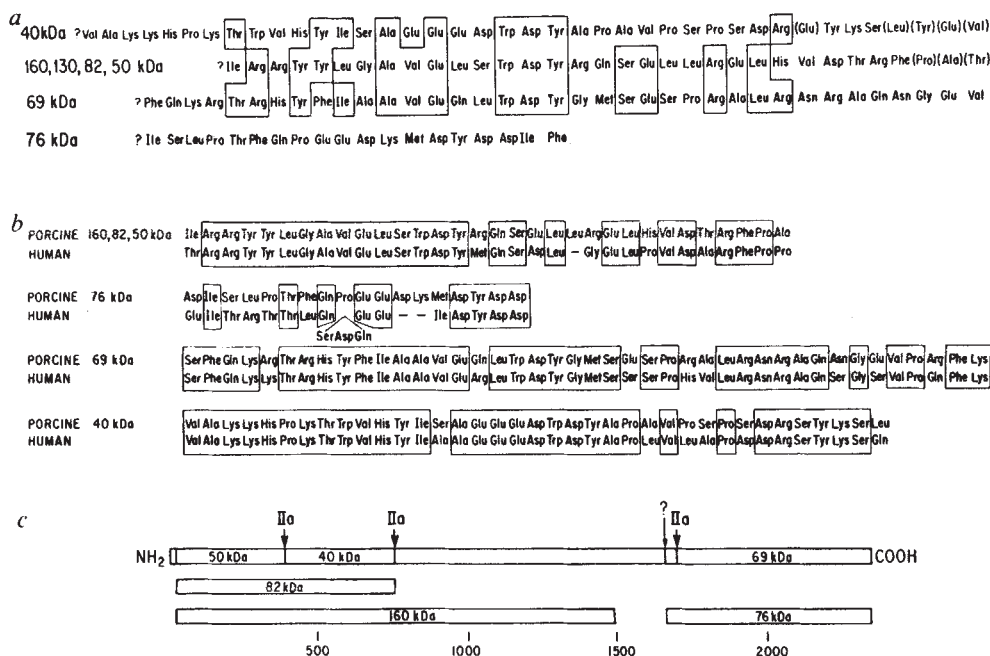
HAEMOPHILIA A is a bleeding disorder caused by deficiency or abnormality of a particular clotting protein, factor VIII:C¹ occurring in about 10-20 males in every 100,000. Afflicted individuals suffer episodes of uncontrolled bleeding and are treated currently with concentrates rich in factor VIII:C derived from human plasma. The available therapy, although reasonably effective, is very costly and is associated with a finite risk of infections. We report here significant progress in the use of recombinant DNA technology to provide pure human factor VIII:C as an alternative treatment for haemophiliacs.

Blood clotting begins with injury to a blood vessel. The damaged vessel wall causes adherence and accumulation of platelets activating the plasma proteins which initiate the coagulation process. Sequential activation, via specific proteolytic cleavages and conformational changes, of a series of proteins comprising the coagulation cascade eventually leads to deposi-

tion of insoluble fibrin which, together with aggregated platelets, curtails the escape of blood through the damaged vessel wall. Factor VIII:C is a large plasma glycoprotein that functions in the blood coagulation cascade as the cofactor for the factor IXa-dependent activation of factor X. It can be activated proteolytically by a variety of coagulation enzymes including thrombin².

In order to provide factor VIII:C for treatment of haemophiliacs we cloned a full-length cDNA. A major obstacle to the cloning effort was the large size of the protein, estimated to be at least M_r 250,000. Purification of factor VIII:C from plasma³ is made difficult by its low abundance, its extreme sensitivity to degradation by serum proteases and its tight association with polymeric forms of the more abundant protein, von Willebrand factor. Fass *et al.*⁴ have described a purification procedure for porcine factor VIII:C using monoclonal antibody

Fig. 1 a, b; N-terminal amino acid sequence analysis of bovine thrombin cleavage and natural cleavage fragments of porcine factor VIII:C. **c,** The positions of fragments in the factor VIII:C precursor in diagrammatic form. The molecular masses are the apparent porcine values deduced from SDS gel electrophoresis rather than those calculated from the human amino acid sequence. Map lengths are indicated in amino acid residues to correlate with Fig. 4. In **a**, N-terminal sequences were determined for the indicated fragments of porcine factor VIII:C. The first residue in each sequence could not be assigned and is indicated as ? In **b**, the porcine amino acid sequences are compared with the corresponding sequences in human factor VIII:C predicted from the nucleotide sequence of the cDNA clone (Fig. 4). Boxed areas, identity between sequences. Thrombin cleavage fragments of porcine factor VIII:C were purified by SDS polyacrylamide gel electrophoresis and located in the unfixed gel using a small portion of ^{125}I -labelled porcine factor VIII:C as a tracer. After electrophoretic elution from the gel matrix as described previously²⁸, the polypeptide sample was subjected to N-terminal amino acid sequence analysis using an Applied Biosystems gas-phase sequencer¹¹.



immunoaffinity chromatography, obtaining highly purified polypeptides of M_r 160,000, 130,000 and 76,000, similar in size to human factor VIII:C polypeptides⁵. Although it has been generally supposed that factor VIII:C, like its homologue factor V, is synthesized as a single chain macromolecular precursor which is then cleaved to yield the fragments found in purified preparations, the results presented here are the first conclusive demonstration that this supposition is correct⁶.

Additional obstacles to the isolation of a factor VIII:C cDNA clone were the lack of a cell line which made factor VIII:C and uncertainty about its site of synthesis *in vivo*⁷⁻⁹. Microsequencing of highly purified porcine factor VIII:C yielded sufficient amino acid sequence to construct oligonucleotide probes. We used these probes to identify a partial clone by direct screening of porcine genomic DNA libraries. The clone provided a probe for identification by cross-species homology of a corresponding human genomic clone. Using this probe, we determined that human liver is a source of factor VIII:C mRNA and obtained subsequently from such tissue the entire 9 kilobase (kb) cDNA sequence as a set of overlapping clones. Reconstruction of these clones into one continuous sequence within an expression vector provided a functional full-length cDNA able to express active protein after transfection into mammalian cells.

Porcine factor VIII:C

Treatment of highly purified porcine factor VIII:C with bovine thrombin (factor IIa) causes an initial activation of up to 60-fold⁴ concomitant with loss of the M_r ~160,000 polypeptide, transient appearance of a M_r 130,000 and the appearance of a new chain of M_r 82,000 (Fig. 1c). As the digestion proceeds, the M_r 76,000 polypeptide generates a fragment of M_r 69,000 and the M_r 82,000 polypeptide is cleaved, apparently to yield polypeptides of M_r ~50,000 and 40,000.

Our sequence analysis was performed using factor VIII:C purified from ~10 l of porcine plasma collected from about ten hogs with precautions to minimize proteolysis during collection and subsequent processing¹⁰. We prepared 0.1–1.5 nmol of each of the M_r 160,000, 130,000 and 76,000 chains originally present in purified porcine factor VIII:C and the 82,000, 69,000,

50,000 and 40,000 thrombin cleavage fragments. N-terminal amino acid sequence analysis using the microsequencing procedures described previously¹¹ yielded the data presented in Fig. 1a, b. The M_r 160,000, 130,000, 82,000 and 50,000 chains have the same N-terminal amino acid sequence for at least thirty residues, whereas the 69,000 and 40,000 fragments have distinct N-terminal sequences. There is, however, striking homology between the 69,000, 40,000 and the 160,000 polypeptides and their fragments. The M_r 76,000 fragment shows no N-terminal sequence homology with the N-terminal sequences of any of the other fragments. Comparison of these sequences with partial sequence data for bovine factor V has led to a hypothetical model for the structural relationship between the putative single chain factor VIII:C and its cleavage products⁶ similar to that shown diagrammatically in Fig. 1c.

Genomic cloning

As the tissue source of mRNA for factor VIII:C was unknown, we tried the novel approach of directly screening genomic libraries with oligonucleotide mixtures to obtain a unique length of porcine factor VIII:C DNA. We used two probes, one a pool of long oligomers containing few sequences and the other a pool of shorter oligomers containing all possible sequences. We identified residues His8–Met22 of the M_r 69,000 chain (see Fig. 1) as containing the least codon ambiguity. The set of 45-mers described in the legend to Fig. 2 was selected on the basis of eukaryotic codon usage¹², the relative stability of G·T versus A·C mismatches¹³ and the infrequency of the dinucleotide CpG in eukaryotic genes. Jaye *et al.*¹⁴ used a similar strategy to generate long oligonucleotides for screening a cDNA library, but they did not attempt genomic screening. A set of 15-mers containing all possible sequences that could encode the Trp18–Met22 region of the M_r 69,000 polypeptide served as the secondary probe.

We constructed a genomic library from *Bam*HI digested female porcine DNA and screened it with the radiolabelled 45-mers under relatively nonstringent conditions (see Fig. 2 legend). Rescreening the library with the 15-mers yielded four clones out of 4×10^5 recombinants which hybridized strongly to

both the 45-mers and the 15-mers. The four bacteriophages were isolated and, on hybridization to additional oligonucleotides derived from the amino acid sequence of the M_r 69,000 polypeptide, one contained a 6.6 kilobase (kb) *Bam*HI fragment that hybridized to all probes (Fig. 2). Sequence analysis of DNA fragments subcloned into M13 confirmed that we had isolated a recombinant containing DNA, encoding a portion of the M_r 69,000 polypeptide. The 45-mer had 11% mismatch (5 of 45 nucleotide positions) but still hybridized strongly to the porcine genomic clone.

To obtain a human factor VIII:C genomic clone, the 6.6 kb *Bam*HI fragment from the porcine genomic clone was used to screen a human *Hae*III/*Alu*I library¹⁵. From 8×10^5 screened, a single recombinant phage containing an insert of about 16 kb was isolated (Fig. 2). The region of this clone containing sequence homologous to the porcine M_r 69,000 fragment, identified by hybridization to the 45-mers, was subcloned into M13 and its sequence determined. The sequence shows strong (>80%) homology to the porcine factor VIII:C sequence at both the nucleotide and derived amino acid level (see Fig. 1). Additionally, an exon/intron boundary occurs at precisely the same position in both porcine and human genomic DNAs.

To confirm further that we had cloned a segment of the human factor VIII:C gene, we demonstrated linkage to the X chromosome by comparing the Southern blot band intensities between human male diploid DNA and XXXXY DNA (human lymphoblastoid line GM 1202A from the NIGMS mutant cell repository). Several unique DNA fragments from the human genomic clone were used as probes. Each hybridized to a single human fragment which was appropriately overrepresented in the 4X DNA when compared with human albumin DNA internal controls.

The entire human factor VIII:C gene was isolated subsequently as overlapping clones spanning ~200 kb. We thus constructed a genomic library from the human XXXXY cell line and screened sequentially with unique probes from the ends of previously isolated recombinants. The factor VIII:C gene spans more than 180 kb; most of its exons are small (150–250 base pairs) with the exception of the 3.1 kb exon (Fig. 2) and the 1.9 kb 3'-most exon, which includes 1.8 kb of non-coding sequence.

Table 1 Expression of factor VIII:C cDNA in COS-7 cells

	Chromogenic activity (mU ml ⁻¹)	1 stage APTT assay (mU ml ⁻¹)
No DNA	<0.05	<<
pCVSVL	<0.05	<<
pCVSVL-VIII	7	10.0
+ thrombin	ND	50
+ α VIII IgG	<0.05	ND

The full-length cDNA clone was excised from pSP64-VIII with *Sal*I and inserted into the *Pst*I site in pCVSVL¹⁸ by addition to the vector of synthetic oligonucleotide adapters containing *Sal*I cohesive ends. The resulting plasmid containing the factor VIII:C cDNA in the correct orientation for transcription from the adenovirus major late promoter was designated pCVSVL-VIII. Plasmid DNA (8 μ g) was transfected into COS-1 monkey cells²³ in a 10 cm dish by the DEAE-dextran transfection protocol²⁴ with the addition of chloroquin treatment²⁵. The transfected cells were incubated for 36 h, then rinsed and fed with 4 ml serum-free media. Samples were taken for assay 24 h later. Factor VIII:C activity was determined by the Kabi Coatest factor VIII:C method modified to afford a sensitivity better than 0.05 mU ml⁻¹ and by the one-stage activated partial thromboplastin time (APTT) coagulation assay²⁶ using factor VIII:C-deficient plasma. For thrombin activation, samples were pretreated 1–10 min with 0.2 units ml⁻¹ thrombin at room temperature. The peak of activity (value shown) was reached after 4 min and then gradually declined. Inhibition with anti-factor VIII:C immunoglobulin was measured by the standard Bethesda protocol²⁷. Isolated immunoglobulin from serum of a patient with high titre factor VIII:C inhibitor was prepared by ion exchange and protein A Sepharose chromatography¹⁰. ND, not determined; <<, not detectable.

cDNA cloning

Our original human factor VIII:C genomic clone provided a probe with which to identify a useful tissue source for human factor VIII:C mRNA. A cell line secreting factor VIII:C is not available, but it had been reported that the protein was synthesized in liver and spleen^{7–9}. RNA blot analysis, initially with porcine material, confirmed the presence of very low mRNA levels in these tissues. Northern blot analysis of human liver mRNA, using a highly radioactive single-stranded probe of ~800 nucleotides of human factor VIII:C exon DNA (see Fig. 3 legend), demonstrated a single mRNA of ~9 kb (Fig. 3A, lanes a, b). A messenger of this size could encode a single-chain factor VIII:C precursor of M_r 250,000–300,000. In other experiments, the same size mRNA was detected in placenta but was absent from peripheral blood lymphocytes. All of these observations are consistent with the suggestion that factor VIII:C mRNA is synthesized by endothelial cells¹⁶.

We demonstrate in Fig. 3A (lane c) that the level of factor VIII:C mRNA in human liver is far lower than that of coagulation factor IX mRNA. Quantitative RNA dot blot hybridizations indicated that the amount of factor VIII:C mRNA was 20–40 times lower than factor IX mRNA. We had found previously that 1 in 5,000 recombinant phage in a liver cDNA library contains factor IX sequences. Thus we estimate that libraries of more than 200,000 recombinants would be required to isolate a single factor VIII:C clone.

Several large cDNA libraries were constructed in λ Charon 21A and GT10 (see Fig. 3 legend), using as a first strand primer either oligo(dT) or a unique 38-mer deduced from the sequence of the large human exon in Fig. 2. Ten factor VIII:C cDNA clones were isolated after screening ~2,000,000 recombinants

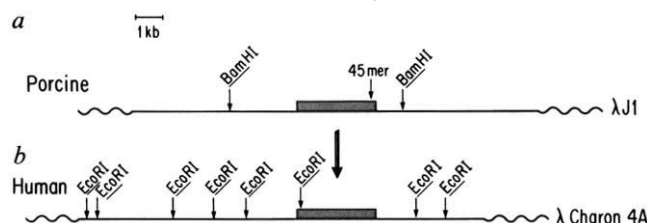


Fig. 2 a, Porcine and b, human factor VIII:C genomic clones. Homologous 3.1 kb exons are indicated by shaded boxes.

Methods: A porcine genomic library was constructed from *Bam*HI digested female porcine liver DNA in the bacteriophage vector λ J1 (a derivative of L47.1) using established procedures. Recombinants (4×10^5) were screened using a modification of the procedure of Woo²⁹. A set of 45-mers containing the sequences 5'-d(CATGCCATA⁺TCCCACAGCTG⁺TCCACAGCAGCAAT⁺AGTAGTG) was synthesized on the Applied Biosystems Model 380A instrument. Filters were hybridized with 5×10^6 c.p.m. ml⁻¹ (0.5 pmole ml⁻¹) of ³²P-labelled 45-mers for 12 h at 45 °C in 5X SSC, 5X Denhardt's, 0.1% SDS, 100 μ g ml⁻¹ denatured salmon sperm DNA. Filters were washed at 50 °C in 5X SSC, 0.1% SDS, air dried and subjected to autoradiography using Kodak XAR film. After autoradiography, filters were denatured in 0.1M NaOH, neutralized and rehybridized with ³²P-labelled 15-mers, sequence 5'-d(CATNCC⁺TA⁺TCCCA). Hybridization was performed as described above at 37 °C. Recombinants hybridizing to both oligomers were plaque purified and phage DNA prepared. Phage DNA prepared from one recombinant (PB34) contained a 6.6 kb *Bam*HI fragment which hybridized strongly to both sets of oligonucleotides in Southern blot analysis. The 6.6 kb *Bam*HI insert was purified from an agarose gel and ³²P-labelled by nick translation. Recombinants (8×10^5) from the human *Hae*III/*Alu*I genomic library of Lawn *et al.*¹⁵ were hybridized with 10⁶ c.p.m. ml⁻¹ of the nick-translated porcine probe using the same conditions described for the 45-mer hybridization. A single recombinant phage, HH25, was thus isolated. *Sau*3A fragments of the insert from HH25 were subcloned into M13 and a phage which hybridized to the radio-labelled 45-mer was identified. The sequence of the *Sau*3A fragment in this subclone proved that HH25 contained DNA from the human factor VIII:C gene. The entire exon (shaded box) was subsequently sequenced. A 0.8 kb *Bam*HI fragment spanning the 3' exon/intron junction was subcloned as the probe to screen for a tissue source of factor VIII:C mRNA (see Fig. 3).

from the oligo(dT) libraries; 55 additional cDNA clones were obtained from among 1,000,000 recombinants from the 38-mer primed library. Four of the several clones characterized in detail are illustrated in Fig. 3B together spanning 9,000 base pairs (bp), sufficient to encode the entire mRNA detected by Northern

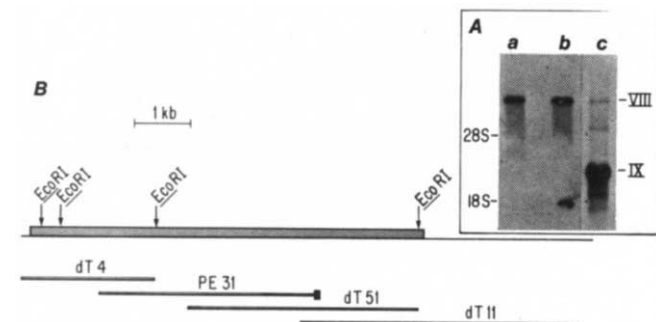


Fig. 3 A, Northern blot analysis to locate a tissue source for human factor VIII:C mRNA. Polyadenylated RNA (5 µg) extracted³⁰ from adult liver (lane b) or fetal liver (lanes a, c) was fractionated on a 0.8% agarose formaldehyde gel and transferred to nitrocellulose as described³¹. The filter in lanes a, b was hybridized with 5×10^6 c.p.m. of a 32 P-labelled factor VIII:C single-stranded probe prepared by primed synthesis from an M13 template which contained an 0.8 kb base *Bam*HI fragment of HH25 (see Fig. 2 legend)^{32,33}. The filter in lane c was probed in the same manner except 5×10^6 c.p.m. of a human factor IX probe of identical specific activity was used in addition. The exposure time for lanes a, b was longer than that for lane c and the RNA was from a different preparation. B, restriction map (*Eco*RI) of the human factor VIII:C cDNA sequence showing the single long open reading frame as a shaded box. An illustrative set of four overlapping bacteriophage λ clones are shown. Those designated dT4, dT51 and dT11 were from oligo(dT) primed λCharon 21A libraries. Clone PE31 was from a library in λGT10 primed with a unique 38-mer (solid rectangle). The cDNA sequence comprises 150 bases of 5'-untranslated region, an open reading frame of 7,053 bases and 1,806 bases of 3'-untranslated region.

Methods: Oligo(dT) primed double-stranded cDNA was synthesized from 10 µg polyadenylated fetal liver RNA as described³⁴. 200 pmol of a unique complementary 38-mer (nucleotides 5,308–5,345) and 10 µg polyadenylated fetal liver mRNA were denatured in 10 mM CH_3HgOH . The reaction was made 140 mM in 2-mercaptoethanol, 700 mM KCl, 1 mM EDTA, 20 mM Tris-HCl (pH 8.3 at 42°C), 1 unit μl^{-1} of RNasin (Promega-Biotec) and incubated at 50°C for 2 min and at 42°C for 2 min. First and second strand synthesis was then performed as for the oligo(dT) primed cDNA³⁴. The second strand reaction was terminated by the addition of EDTA to 20 mM. The *Eco*RI sites were methylated, after the addition of S-adenosylmethionine to 50 µM and 40 units of *Eco*RI methylase, by incubation at 37°C for 1 h. These reactions were terminated by phenol-chloroform extraction and chromatographed using Sephadex G-50 equilibrated in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 100 mM NaCl. DNA in the excluded volume was pooled and precipitated with ethanol. The double-stranded cDNA was blunted in 50 µl containing 50 mM Tris pH 8.3, 10 mM MgCl_2 , 10 mM 2-mercaptoethanol, 50 mM NaCl, 50 µM of deoxynucleotide triphosphate, 100 $\mu\text{g ml}^{-1}$ ovalbumin and 5 units of T4 polymerase (PL Biochemicals). The reaction was incubated at 37°C for 30 min and then terminated by phenol-chloroform extraction. Nucleic acids were recovered by ethanol precipitation with 2 M ammonium acetate. Kinased *Eco*RI linkers (NE. Biolabs; 300 ng) were then added to the blunted double stranded cDNA by ligation overnight at 16°C in a final volume of 30 µl. The reaction was diluted to 200 µl in *Eco*RI digestion buffer and digested with 300 units of *Eco*RI (NE. Biolabs) for 2 h at 37°C. EDTA to 15 mM was added and the reaction extracted with phenol-chloroform and chromatographed over Sepharose CL-4B (Pharmacia) equilibrated with 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 100 mM NaCl. cDNA in the void volume was collected and precipitated by ethanol in the presence of 10 µg yeast tRNA carrier. cDNA was redissolved in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA and ligated to *Eco*RI cleaved, phosphatased, λCharon 21A DNA for the oligo(dT) primed cDNA and to λGT10 DNA for the primer extended cDNA. The libraries were packaged, plated and initially screened using a 32 P-labelled 53-mer (nucleotides 5,158–5,210) permitting the isolation of both PE31 and dT51. Recombinant dT4 was identified with a 32 P-labelled pool of 64 56-mers containing the sequences; 5'd(TCCCAATCCTCTTCTTCA⁺GC⁺GA⁺A-TATAGTG⁺ACCCA⁺GTCTT⁺GGGTGCTTCTT) (using the same criteria as was used for the construction of the 45-mer pools) made to residues (Lys3-Asp21) of the M_{40,000} polypeptide fragment of porcine VIII:C. Recombinant dT11 was isolated using a nick translated 3' end *Pst*I/*Eco*RI fragment (nucleotides 5,902–7,127) of dT51. Numerous other cDNA clones were isolated by similar methods; some were also characterised in detail, sequenced and used in the construction of the full-length cDNA clone. Selected cDNA and genomic exon fragments were ligated to one another two at a time, subcloned and ultimately inserted in the *Sal*I site of pSP64 after the addition of a synthetic *Sal*I site at nucleotide 145. The resulting recombinant, pSP64-VIII, contains the factor VIII:C cDNA from just before the initiator ATG (at nucleotide 150–152) to nucleotide 7,422 (some 200 bp downstream from the terminator at nucleotide 7,201).

blot analysis. (The unique 9,009 bp cDNA sequence we obtained after sequencing numerous overlapping fragments of cDNA and genomic DNA has been deposited in the NIH US Nucleic Acid Sequence Databank and is available through GENBANK.) The sequence comprises a 5'-noncoding region of 150 nucleotides, a single long open reading frame of 7,053 nucleotides (2,351 codons) and 1,806 nucleotides of the 3'-noncoding region. The translated sequence, in a single-letter amino acid code, is presented as Fig. 4.

A full-length cDNA clone was assembled (see Fig. 3 legend) and inserted into the *Sal*I site of the polylinker in plasmid pSP64 (Promega Biotec). We thus obtained a recombinant, pSP64-VIII, with a single continuous factor VIII:C coding sequence that could be readily excised by *Sal*I digestion for insertion into a variety of expression vectors.

Expression and structure

To confirm that the factor VIII:C cDNA indeed encodes active factor VIII, we introduced the full-length cDNA into the mammalian expression vector pCVSVL^{17,18}. Plasmid pCVSVL-VIII contains the cDNA in the proper orientation to be expressed from the adenovirus major late promoter. Medium collected 60 h after transfection of COS-1 cells with this plasmid contained readily detectable levels of factor VIII:C activity (Table 1). No activity was detected in controls. The activity was inhibited by antibody to human factor VIII:C and was activated fivefold by thrombin.

The sequence of the cDNA clone comprises 2,351 amino acids from the initiator Met to the terminating Tyr residue and begins with a secretory leader peptide sequence of 19 amino acids. Regions of homology with the previously determined amino acid sequences of porcine factor VIII:C fragments can be located (Fig. 1), confirming the model for the structure of porcine factor VIII:C of Fass *et al.*⁶ and predicting the mature N-terminus and thrombin cleavage sites of the human molecule.

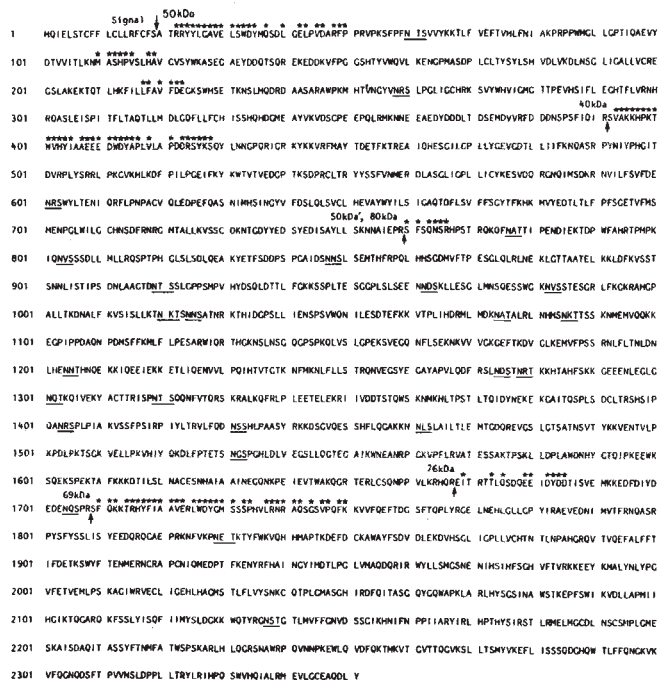


Fig. 4 Amino acid sequence of the human factor VIII:C precursor, predicted from the sequences of cDNA and genomic clones. Arrows indicate (1) a presumed signal peptidase cleavage site at residues 19–20 (Ser-Ala); (2) thrombin cleavage sites at residues 391–392, 759–760 and 1,708–1,709 (all Arg-Ser); (3) a possible thrombin cleavage site at residues 1,667–1,668 (Arg-Glu). Before each cleavage site the apparent M_r of the corresponding porcine factor VIII:C fragment C-terminal to the cleavage site⁶ is given. * Amino acid residue identity with porcine factor VIII:C amino acid sequence determined from a variety of proteolytic and chemical cleavage fragments. Potential N-glycosylation sequences (Asn-X-Ser/Thr) are underlined.

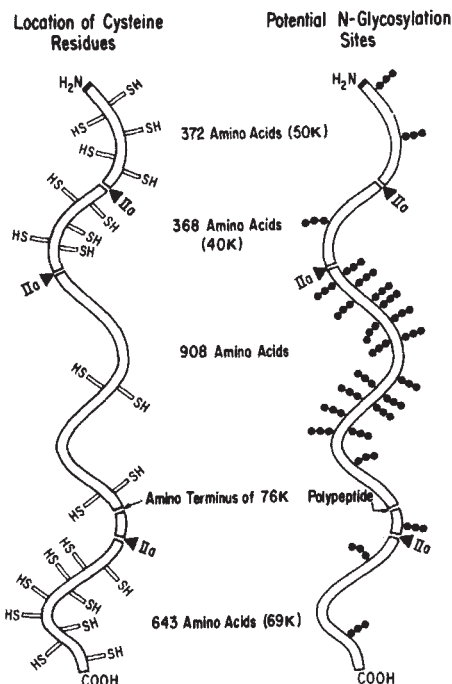


Fig. 5 Structural features of human factor VIII:C, showing schematically the location of cysteine residues and potential N-glycosylation sites (Asn-X-Ser/Thr) with respect to thrombin cleavage domains. The number of amino acid residues in each domain is given as well as the apparent molecular weights of the corresponding regions of porcine VIII:C.

The location of cysteine residues and potential N-glycosylation sites (Asn-X-Ser/Thr) is shown schematically in Fig. 5. Their exact numerical position within the human factor VIII:C precursor molecule is shown in Fig. 4. The cysteine residues appear to be clustered in the M_r 50,000, 40,000 and 69,000 regions. There are no disulphide bridges between these regions in porcine factor VIII:C, where the same pattern of polypeptides is observed on SDS gels in either reducing or non-reducing conditions.

In contrast to the location of cysteine residues, many of the potential N-glycosylation sites appear between the M_r 40,000 and 69,000 regions (Fig. 5). Distinct bands corresponding to this domain of 908 amino acids, or cleavage fragments derived from it, cannot be readily identified on silver stained SDS gels. We have obtained amino acid sequence data from minor components in purified porcine factor VIII:C preparations which are present in the M_r 80,000 and 50,000 regions of SDS gels. The N-terminal sequences of these polypeptides are identical; furthermore, a similar amino acid sequence is found in the predicted human factor VIII:C precursor amino acid sequence shown in Fig. 4 (residues 760-766).

A striking feature of human factor VIII:C is the presence of repeated blocks of amino acid homology (Fig. 6); the three most extensively repeating patterns of homology are within the M_r 50,000 and 69,000 fragments (blocks A, B, C, D, E and A', B', C', D', E'); within the M_r 50,000 and 40,000 fragments (blocks A, F, G, H, I and A', F', G', H', I') and a M_r 69,000 fragment internal repetition (blocks J, K, L and J', K', L'). Another set of repeat blocks (A', M and A'', M') is suggestive of a vestige of a M_r 40,000/69,000 duplication. Not only is there a strong M to M' homology, but also the A to M spacing is very similar in both regions. A block of amino acids (N) in the 40,000 region is repeated (N') in the region between the M_r 40,000 and 76,000 regions.

We analysed the precursor human factor VIII:C amino acid sequence by the method of Hopp and Woods¹⁹ for the determination of relative hydrophobicity. As observed for other secreted proteins, the factor VIII:C precursor contains a hydrophobic

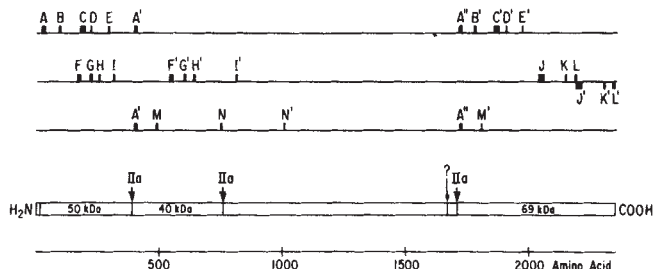


Fig. 6 Amino acid sequence homology within human factor VIII:C. A computer search for internal amino acid sequence homology was performed using the PEP program in the IntelliGenetics software system. Blocks of homology are labelled alphabetically on the linear maps, with their location and relative size indicated. Matching blocks of homology bear the same alphabetical symbol plus a prime or double prime. The actual amino acid sequences within these blocks and their numerical position with the entire precursor molecule can be located in Fig. 4. A(residues 21-35); A'(400-414); A''(1,714-1,728); B(89-101); B'(1,776-1,788); C(177-198); C'(1,856-1,877); D(211-227); D'(1,902-1,908); E(292-298); E'(1,979-1,985); F(164-179); F'(539-554); G(217-227); G'(594-604); H(250-261); H'(633-644); I(318-324); I'(809-815); J(2,038-2,059); J'(2,191-2,212); K(2,144-2,151); K'(2,301-2,308); L(2,181-2,188); L'(2,338-2,345); M(486-492); M'(1,796-1,802); N(749-755); N'(1,002-1,008).

leader sequence (residues 1-19). Two other extensive regions of hydrophobicity are also predicted by this analysis; first, 27 amino acids from Ser308 to His334 and second, from Ser630 to Phe677. The hydrophobic nature of these regions is substantiated by the method of Kyte and Doolittle²⁰. The role of these hydrophobic regions in factor VIII:C secretion and/or activity is unknown.

Homology with other plasma proteins

We have reported previously²¹ amino acid sequence homologies among certain regions in factor VIII:C, factor V, and ceruloplasmin. These homologies were identified by comparison of the N-terminal amino acid sequences of thrombin cleavage fragments of bovine factor V and porcine factor VIII:C with amino acid sequence deduced from a genomic clone of porcine factor VIII:C and with the published amino acid sequence of human ceruloplasmin²². The deduced amino acid sequence of human factor VIII:C described here (Fig. 4) contains regions of homology with the known amino acid sequences of bovine factor V and extensive homology with the entire human ceruloplasmin molecule, supporting the suggestion that these three proteins are related evolutionarily. Note the homology between blocks A, A' and A'' of human factor VIII:C (Fig. 6) and residues 2-16, 353-366 and 712-726 respectively of human ceruloplasmin²², as these homology blocks apparently mark the boundary of a primordial gene encoding approximately 350 amino acids. Similarly, block A is homologous with residues 3-17 of the M_r 94,000 amino terminal thrombin cleavage fragment of bovine factor V and block A'' is homologous with residues 6-20 of the C-terminal M_r 74,000 thrombin cleavage fragment of factor V²¹. At present, the significance of repeating homologous units in factor VIII:C, factor V and ceruloplasmin is unclear, but they presumably reflect duplications and triplications of some smaller ancestral gene.

We thank Louise Wasley and John Brown for technical assistance; Ed Fritsch and Steve Clark for support and helpful discussions; Joyce Lauer and Katherine Smith for help with the manuscript and Martha Richardson, Marybeth Erker and Cathy Beechiner for word processing. Finally, we thank both Robert Kamen and Gabriel Schmergel for their encouragement throughout the project. This work was supported by Baxter Travenol Inc., the Mayo Foundation and grant HL-17430 to D.N.F.

Received 31 August; accepted 2 October 1984.

- Hoyer, L. W. *Blood* **58**, 1-12 (1981).
- Rapaport, S. I., Schiffman, S. & Patch, M. J. *Blood* **21**, 221-235 (1963).
- Vehar, G. A. & Davie, E. W. *Biochemistry* **19**, 410-416 (1980).
- Fass, D. N., Knutson, G. J. & Katzman, J. A. *Blood* **59**, 594-600 (1982).
- Fulcher, C. A., Gardiner, J. E., Griffin, J. H. & Zimmerman, T. A. *Blood* **63**, 486-489 (1984).
- Fass, D. N., Hewick, R. M., Knutson, G. J., Nesheim, M. E. & Mann, K. G. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Owen, C. A. Jr, Bowie, W. E. J. & Fass, D. N. *Br. J. Haemat.* **43**, 307-315 (1979).
- Libre, E. P., Cowen, D. H., Watkins, S. P. & Shulman, N. R. *Blood* **31**, 358-368 (1968).
- Marchioro, T. L., Houghie, C., Ragde, H., Epstein, R. B. & Thomas, E. D. *Science* **163**, 188-190 (1969).
- Knutson, G. J. & Fass, D. N. *Blood* **59**, 615-624 (1982).
- Hewick, R. M., Hunkapiller, M. E., Hood, L. E. & Dreyer, W. J. *J. biol. Chem.* **256**, 7990-7997 (1982).
- Strehler, B. & North, D. *Mech. Ageing Dev.* **18**, 285-313 (1982).
- Agarwal, K. L., Brunstedt, J. & Noyes, B. E. *J. biol. Chem.* **256**, 1023-1028 (1981).
- Jaye, M. *et al. Nucleic Acids Res.* **11**, 2325-2335 (1983).
- Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15**, 1157-1174 (1978).

- Stel, H. V., vanderKwaast, Th. H. & Veerman, E. C. I. *Nature* **303**, 530-532 (1983).
- Kaufman, R. J. & Sharp, P. A. *Molec. Cell Biol.* **2**, 1304-1319 (1982).
- Clark, S. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 2541-2547 (1984).
- Hopp, T. P. & Woods, K. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3824-3828 (1981).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105-132 (1982).
- Church, W. R. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Takahashi, N., Ortel, T. L. & Putnam, F. W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 390-394 (1984).
- Gluzman, Y. *Cell* **23**, 175-182 (1981).
- Sompayrac, L. M. & Dana, K. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7575-7578 (1981).
- Luthman, H. & Magnusson, G. *Nucleic Acids Res.* **11**, 1295-1308 (1983).
- Lee, M. L., Maglalang, E. A. & Kingdon, H. S. *Thromb. Res.* **30**, 511-519 (1983).
- Kasper, C. K. *Thromb. Diath. Haemorrh.* **34**, 869-872 (1975).
- Hunkapiller, M. W., Lujan, E., Ostrander, F. & Hood, L. E. *Meth. Enzym.* **91**, 227-236 (1983).
- Woo, S. L. C. *Meth. Enzym.* **68**, 389-395 (1979).
- Hastie, N. D., Held, W. & Toole, J. J. *Cell* **17**, 449-457 (1979).
- Derman, E. *et al. Cell* **23**, 731-739 (1981).
- O'Hare, K., Levis, R. & Rubin, G. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6917-6921 (1983).
- Shaw, P. H., Held, W. A. & Hastie, N. D. *Cell* **32**, 755-761 (1983).
- Gubler, U. & Hoffman, B. J. *Gene* **25**, 163-269 (1983).

LETTERS TO NATURE

Cosmic-ray emission from Cygnus X-3

N. A. Porter

Physics Department, University College, Dublin 4, Ireland

Recent observations^{1,2} of the X-ray binary source Cygnus X-3 have revealed a flux of ultra-high energy γ rays. Other observations and subsequent models suggest that Cyg X-3 may be a dominant source of galactic high-energy cosmic rays. Here I investigate whether it is possible to escape this conclusion with a model in which most of the energy of accelerated particles at the source is emitted as curvature radiation (that is, radiation emitted by particles travelling on curved paths in a magnetic field so strong that synchrotron radiation is suppressed). I conclude that such a model is not consistent with the observations.

The ultra-high energy γ -ray flux^{1,2} at 10^{15} – 10^{16} eV from the X-ray binary source Cygnus X-3 has attracted much attention. If the output is in the form of a rotating beam or an interception of a particle beam by a producing medium in the line-of-sight, the estimated¹ emission in the energy range is 6×10^{36} erg s⁻¹. The flux at 10^{11} – 10^{13} eV seen by atmospheric Čerenkov groups⁴⁻⁷ also contributes $\sim 10^{37}$ erg s⁻¹, again assuming a beamed emission. This γ -ray component may be significant at high energies in forming cosmic-ray anisotropies⁸. Because of the difficulty in explaining the ultra-high energy flux in terms of electron emission^{2,9,10}, a proton or ion mechanism, probably involving nuclear collisions and production of π^0 -mesons¹¹, seems necessary. Using nuclear interaction models of the source, it has been shown^{12,13} that more than 10^{40} erg s⁻¹ of high-energy cosmic-ray particles are emitted. The calculation of the production and subsequent absorption of γ rays in a uniform medium gives a maximum efficiency of only $\sim 3\%$ for the conversion of a nuclear spectrum into a γ -ray spectrum of comparable exponent. Interstellar absorption, for the accepted distance of 11.4 kpc (ref. 3), raises the observed flux to $\sim 3 \times 10^{37}$ erg s⁻¹ in the region 10^{15} – 10^{16} eV. These are time-averaged fluxes. In a directed effect, in which an isotropic flux of protons or ions strikes an emitting/absorbing medium¹³ and if the observed duty cycle is 0.1 (in ref. 2 it seems to be as low as 0.025), then we arrive at a heavy-particle emission of $\geq 10^{40}$ erg s⁻¹. However, the total cosmic-ray energy injection rate required for the whole Galaxy is thought^{14,15} to be only 10^{40} – 10^{41} erg s⁻¹. Thus a single Cyg X-3 type source could supply the entire cosmic-ray demand up to particle energies of $\sim 10^{16}$ eV.

Hillas¹³ holds that Cyg X-3 can comfortably supply the ultra-high energy galactic cosmic rays in the energy range 10^{15} – 10^{18} eV, but probably does not contribute substantially to the lower energy flux. Does the observed γ -ray flux necessarily imply such large proton fluxes, or are there are mechanisms which could supply the photon component more economically? One possible mechanism, curvature radiation¹⁶, is considered here because it has a very efficient transfer ratio between emitting particles and emitted photons, and because it operates with heavy particles as well as electrons. For the 10^{15} – 10^{16} eV band,

protons or ions are required, but electrons might produce the 10^{11} – 10^{13} eV component. For the 10^{13} – 10^{15} eV interval, for which there is some experimental evidence of emission¹⁷, there could be a combined effect, and indeed the phase of the 4.8-h period emission seems to change between 10^{12} and 10^{15} eV. For either proton or electron emission, however, there is a serious γ -ray absorption problem, arising from pair production in the magnetic field of the pulsar¹⁸. Ignoring this we can calculate, in cgs units, the required power output. For a single particle of unit charge, whether electron or proton, the power output in curvature radiation is:

$$W = 2e^2 c \gamma^4 / 3r^2 \quad (1)$$

With a characteristic γ -ray energy

$$h\nu = 2\pi\hbar c \gamma^3 / r \quad (2)$$

where r is the radius of the orbit, $\gamma = E/mc^2$ for either a proton or an electron, where E is the energy of the proton or electron.

For the 10^{15} -eV γ rays we take $r = 10^8$ cm, for a field line extending from the pulsar pole to the velocity of light cylinder. The frequency $\nu = 10^{30}$ Hz, and from equation (2) we have $\gamma^3 = 3 \times 10^{27}$ or $\gamma = 10^9$. For a proton this gives a particle energy of 10^{18} eV. For orbit radii $< 10^8$ cm, W would be large, resulting in a rapid energy loss, so that the process would not work.

Acceleration to this energy may be electrostatic¹⁰. The field ϵ has a maximum value^{14,15} given by $\epsilon = 10^{-14.2} B r^{3/2} P^{-3/2}$ V cm⁻¹ where B is the magnetic field and P the period; at 10^{12} G and 3 ms respectively we obtain a field of $\sim 2 \times 10^{13}$ V cm⁻¹ near the pole of the pulsar. The distance to the velocity-of-light cylinder is only 10^7 cm, but in the absence of short-circuiting effects, it is not difficult to obtain potentials in the region of 10^{20} eV. An alternative possibility is the mechanism of Gunn and Ostriker¹⁹.

The energy output in curvature radiation from a proton of 10^{18} eV is given by equation (1):

$$W = \frac{2}{3} \times \frac{25 \times 10^{-20} \times 3 \times 10^{10} \times 10^{36}}{10^{16}} \\ = 10^{12} \text{ erg s}^{-1} \text{ or } 30 \text{ erg cm}^{-1}$$

This is essentially equal to the electrostatic energy gained, so that there will be equilibrium between emission and acceleration. For a flight distance 10^7 cm to the velocity-of-light cylinder, we have an emission of 3×10^{18} erg at a constant particle energy of 10^{18} eV. To produce 3×10^{37} erg s⁻¹ of γ rays at 10^{15} eV we require 10^{29} particles s⁻¹, each with 10^{18} eV or 10^6 erg, making a total energy flux of $\sim 10^{35}$ erg s⁻¹. We assume that all particles escape on reaching the velocity-of-light cylinder.

Therefore, the proton flux produced might be less by more than two orders of magnitude than the γ -ray flux, and five orders of magnitude less than the 10^{40} erg s⁻¹ required by the nuclear interaction model. A similar calculation for electrons having characteristic energies of 10^{14} eV results in a comparable energy requirement, about 10^{34} erg s⁻¹ of 10^{14} eV electrons producing

10^{37} erg s $^{-1}$ of γ rays at 10^{12} eV. The assumptions made about magnetic and electrostatic fields are, of course, optimistic, but a more accurate calculation would probably not differ from the above by more than two orders of magnitude.

There is, however, a severe difficulty in extracting the γ rays from the system without pair production in the magnetic field. There will probably also be considerable but lesser geometrical difficulties in producing the characteristic one or two narrow γ -ray pulses in the 4.8-cycle.

The condition for pair production is $B_{\perp}E_{\gamma} > 10^{18.6}$, with B in gauss and photon energy in electron volts 16 . For $E_{\gamma} = 10^{15}$ eV the transverse components of B must be $\leq 10^4$ G. With a surface field of 10^{12} G and an orbit radius of 10^8 cm this cannot be achieved, and escape would be possible, if at all, only from a position very close to the velocity-of-light cylinder. The difficulties are not quite so severe for 10^{12} -eV γ rays. The field at 10^7 cm would be $\leq 10^9$ G, and the angle between the photon trajectory and field line could be $< 10^{-2}$, if the orbit radius were increased to 10^9 - 10^{10} cm. This, however, involves increases in electron

energies and fluxes to a point where $\sim 10^{38}$ - 10^{39} erg s $^{-1}$ are needed to produce 10^{37} erg s $^{-1}$ of 10^{12} -eV γ rays. For protons and 10^{15} -eV γ rays it is difficult to see how the model can work without requiring particle fluxes approaching the 10^{40} erg s $^{-1}$ demanded by the nuclear interaction model.

Another model involves a black hole 20 . It is not yet clear whether γ rays of 10^{15} eV can be produced in this. Alternatively, emission by quantum synchrotron radiation from electrons in high magnetic fields is possible, but this seems to demand very artificial geometries: electrostatic acceleration of the particles in a region essentially free from magnetic field, followed by almost discontinuous injection into the region of a high field. Otherwise, dissipation of particle energy by low-energy synchrotron emission seems inevitable. Similar models with protons as the radiating particles seem to require unrealistically high particle energies.

The nuclear interaction model needs to be developed to see whether it can produce the experimentally-required spectra and nuclear constitution observed in cosmic rays.

Received 8 May; accepted 10 September 1984.

- Samorski, M. & Stamm, W. *Astrophys. J. Lett.* **268**, L17-22 (1983).
- Lloyd-Evans, J. *et al. Nature* **305**, 784-787 (1983).
- Cawley, M. F., Gibbs, K. & Weekes, T. C. *Pap. XG4-25, Proc. 18th int. Conf. on Cosmic Rays*, Bangalore (Tata Institute, Bombay, 1983).
- Nesphor, Yu. I. *et al. Astrophys. Space Sci.* **61**, 349-355 (1979).
- Danaher, S., Fegan, D. J., Porter, N. A. & Weekes, T. C. *Nature* **289**, 568-569 (1981).
- Lamb, R. C., Godfrey, C., Wheaton, W. & Tumer, T. *Nature* **296**, 543-544 (1982).
- Gibson, A. I. *et al. Proc. Ootacamund Workshop on Gamma-ray Astronomy*, 97-117 (Tata Institute, Bombay, 1982).
- Wdowczyk, J. & Wolfendale, A. W. *Nature* **305**, 609-610 (1983).
- Porter, N. A. *Nature* **305**, 179-180 (1983).

- Eichler, D. & Vestrand, W. J. *Nature* **307**, 613-614 (1984).
- Stenger, V. J. Preprint, Hawaii Univ. (1984).
- Weekes, T. C. *Am. Phys. Soc. Abstr.*, Washington, (April 1984).
- Hillas, A. M. Preprint, Leeds Univ. (1984).
- Ramakrishnan, A. *Elementary Particles and Cosmic Rays*, Madras (Pergamon, Oxford, 1962).
- Weekes, T. C. *High Energy Astrophysics* (Chapman and Hall, London, 1969).
- Sturrock, P. A. *Astrophys. J.* **164**, 529-556 (1971).
- Morello, C., Navarra, G. & Vernetto, S. *Pap. XG4-13, Proc. 18th int. Conf. on Cosmic Rays*, Bangalore (Tata Institute, Bombay, 1983).
- Erber, T. *Rev. mod. Phys.* **38**, 626-659 (1966).
- Gunn, J. E. & Ostriker, J. P. *Phys. Rev. Lett.* **22**, 728-731 (1969).
- Grindlay, J. E. *Proc. Ootacamund Conf. on Gamma-ray Astronomy*, 178-196 (Tata Institute, Bombay, 1982).

A disk origin for superhumps in Ursae Majoris variables

Keith Horne

Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

As cataclysmic variables, SU Ursae Majoris stars 1,2 are short-period binary systems consisting of a red dwarf star that fills its Roche-lobe and spills material to an accretion disk around a white dwarf. SU UMa systems undergo outbursts (episodes of increased mass transfer in the disk) of two distinct types: normal outbursts that last a few days, and superoutbursts that last up to 2 weeks. Superhumps are periodic $\sim 30\%$ brightness variations of unknown origin seen only during the superoutbursts. The superhump period is several per cent longer than the binary orbital period. Both the accretion disk 3,4 and the red dwarf 5 have been proposed as possible sites for the superhump. Although superhumps occur in both eclipsing and non-eclipsing systems, observations of the eclipsing systems Z Cha and OY Car may help to discriminate between these two classes of models. Whitehurst *et al.* 6 have shown that the superhumps in Z Cha are not deeply eclipsed, and argue that superhumps therefore cannot originate in the accretion disk. In their starspot model, superhumps arise in a bright active region located in the equatorial belt of the red dwarf. By contrast, the light curve synthesis results reported here show explicitly that disk models produce a shallow partial eclipse of the superhump light, compatible with the eclipse depth measurements of Whitehurst *et al.*, provided the superhump source is not concentrated into a small region of the disk. The starspot model is shown to be inconsistent with the observed shapes of superhumps unless radiation emerges from the starspot in a narrow ($\leq 35^\circ$) fan beam with its long axis in the direction of stellar latitude.

Z Cha is depicted in Fig. 1 as viewed from the Earth at three different orbital phases (ϕ_{orb}). Deep eclipses of the optical light from the system are seen at $\phi_{\text{orb}} = 0.0$ (Fig. 1c), when the red dwarf occults the accretion disk, but secondary eclipses are not seen at $\phi_{\text{orb}} = 0.5$ (Fig. 1a), when the disk occults one hemisphere of the red dwarf. Successive eclipses of the accretion disk occur progressively earlier in the superhump cycle, because the super-

hump period is a few per cent longer than the 107-min orbital period. Eclipse light curves thus probe changes in the structure of the accretion disk during the superhump cycle, and can be used to establish the site of the superhump source.

To determine whether the superhumps in Z Cha are eclipsed, Whitehurst *et al.* measured the brightness of Z Cha at moments of minimum light, when much of the accretion disk is occulted by the red dwarf (Fig. 1c). To correct approximately for secular changes in the brightness of Z Cha during the superoutburst, the fluxes at mid-eclipse are first divided by fluxes at nearby phases that are not affected by either the eclipse or the superhump. Complete coverage of the superhump cycle is achieved by combining measurements from three different superoutbursts. Whitehurst *et al.* found that the normalized mid-eclipse fluxes increased by $\sim 25\%$ for eclipses that occur near the superhump peak, and were roughly consistent at all superhump phases with the expected fluxes for an uneclipsed superhump. Thus the superhump source in Z Cha is not deeply eclipsed.

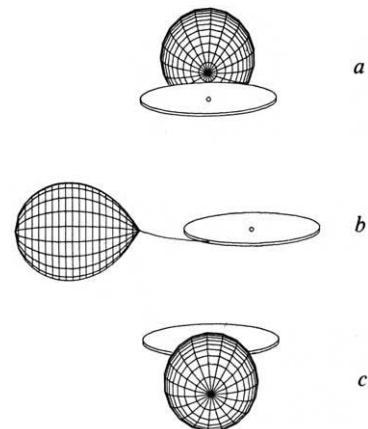


Fig. 1 Schematic views of the cataclysmic variable star Z Cha, at three different binary phases. The red dwarf star is severely distorted by tidal stresses induced by its white dwarf companion. An accretion disk surrounding the white dwarf is fed gas in a stream from the red dwarf originating at the inner lagrangian point. One hemisphere of the red dwarf is largely eclipsed by the disk (a), and a large area of the disk remains visible at mid-eclipse (c).

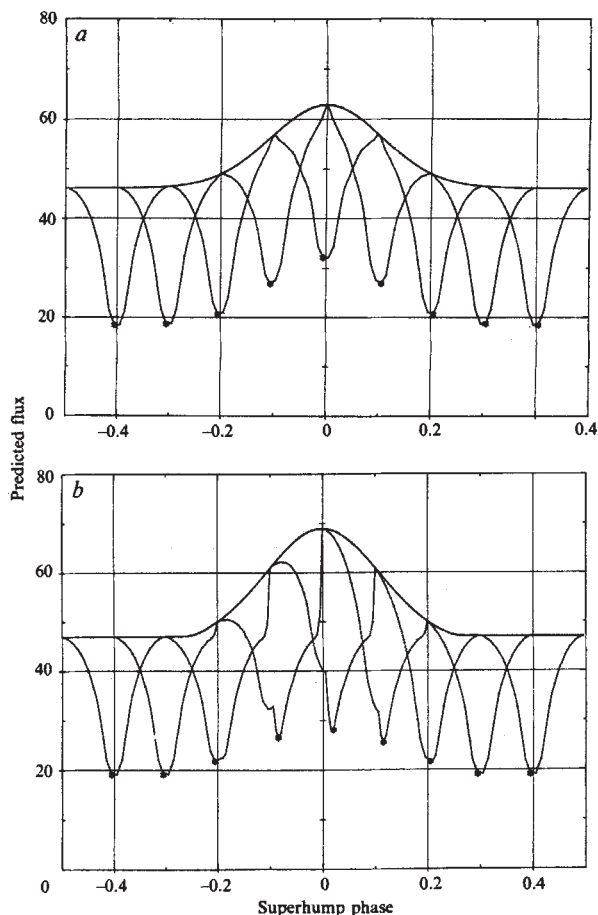


Fig. 2 Synthetic light curves for a disk model of superhumps. The upper envelope, giving the superhump light curve in the absence of eclipses, is produced in the model by modulating the surface brightness in the outer annulus of the accretion disk. Eclipses of the disk at different superhump phases are shown. *a*, A complete full annulus at the outer rim of the disk is modulated. The eclipse shapes in this case remain similar throughout the superhump. *b*, The modulation is confined to a single point on the outer rim of the disk. The distorted eclipses found in this case are not observed, ruling out a compact region on the disk as the site of superhumps. In both cases, the mid-eclipse flux is larger during the superhump because the accretion disk is only partially eclipsed.

The absence of deep eclipses of the superhump light in Z Cha does not preclude a disk origin for superhumps. Although the bright central region of the accretion disk is eclipsed, a large fraction of the disk surface remains visible at minimum light (Fig. 1c). A superhump source that is broadly distributed on the face of the accretion disk therefore undergoes a partial eclipse which may be quite shallow, for example, if the superhump light is distributed around the outer annulus of the disk. Thus, the eclipse depth measurements of Whitehurst *et al.* are compatible with partially-eclipsed superhumps produced in the accretion disk. Smak⁶ independently reaches this same conclusion.

The viability of a disk origin for superhumps is explicitly demonstrated by the synthetic eclipse light curves shown in Fig. 2. Here the uneclipsed superhump light curve is the upper envelope of nine individual eclipse light curves, each centred at a different superhump phase. Fluxes at minimum light, marked by stars in Fig. 2, increase during the superhump in the manner required by the eclipse depth measurements of Whitehurst *et al.* The times of minimum light are slightly advanced during the rise of the superhump, and delayed during the decline. Thus the eclipse phase shift discussed by Whitehurst *et al.* in the context of the starspot model is also produced by a disk model.

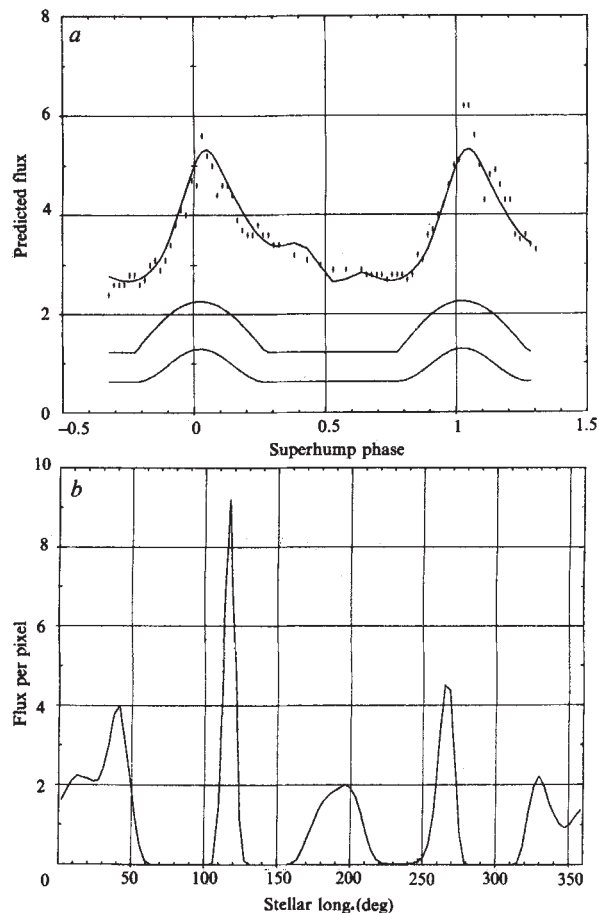


Fig. 3 The synthetic light curve for a rotating spherical star with a strip of non-uniform brightness around the stellar equator is compared in *a* with an observed superhump light curve of Z Cha. The longitudinal brightness distribution of the fitted model is given in *b*. The light curve peak produced by this starspot model is always broader than the observed superhump, because the light curves of point sources on the stellar equator (undarkened and fully limb-darkened cases are shown) are already too wide. Flux is given in arbitrary units.

The qualitative behaviour illustrated by this simulation of superhumps from the accretion disk is not sensitive to errors in the assumed geometry of the Z Cha system, to specific details of the accretion disk model, or to the assumed spatial distribution and temporal development of the superhump light, provided only that the superhump source is broadly distributed on the accretion disk. The eclipse geometry (Fig. 1) is the same as that assumed by Whitehurst *et al.* The accretion disk brightness distribution is a steady-state black-body accretion disk model that roughly reproduces the shape and depth of the observed eclipses of Z Cha during superoutburst. The red dwarf is assumed to contribute no light.

The uneclipsed shape of the superhump can be specified arbitrarily and is therefore not an important feature of the simulation. The superhumps in Fig. 2 are produced by modulating the temperature (T) of the outer annulus of the accretion disk according to $T \propto \exp(-2 \cos \phi_{sh})$, where ϕ_{sh} is the superhump phase. This specific form is motivated by the model of Papaloizou and Pringle³ in which mass transfer from the red dwarf varies exponentially with small changes in the binary separation caused by a slightly eccentric orbit, while the discrepant orbital and superhump periods are explained as a precession of the line of apsides.

If the superhump light curve in the absence of eclipses is known, the observed eclipse light curves give detailed information about the spatial distribution of the superhump source on the accretion disk. The eclipse shapes are not strongly affected

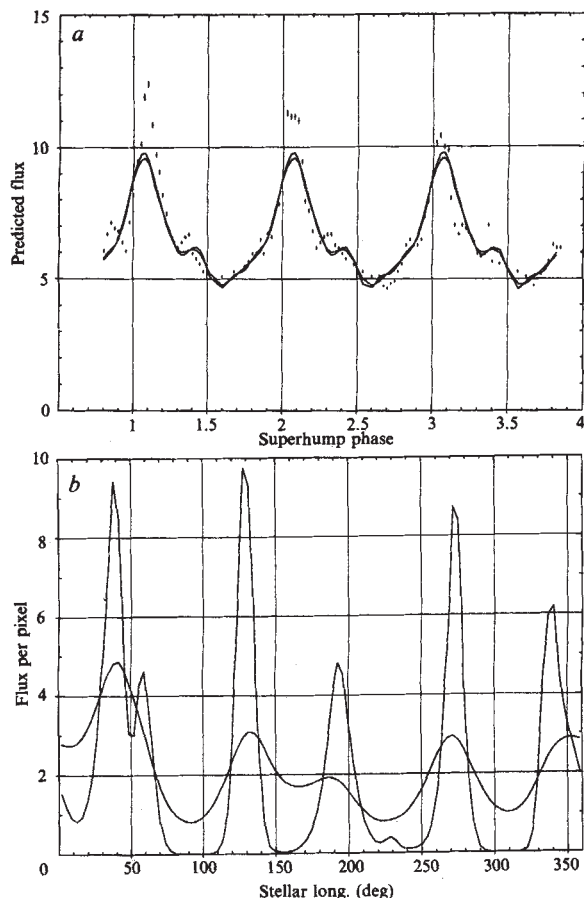


Fig. 4 The starspot model fails to produce the sharply-peaked superhumps observed in OY Car. Two models with different longitude resolution are shown. *a* and *b* are as in Fig. 3.

by the superhump if, as in Fig. 2*a*, the superhump light is uniformly distributed around the outer annulus of the disk. A superhump source confined to a compact region on the accretion disk produces highly distorted eclipses (Fig. 2*b*). Because such distortions are not observed, the main source of superhump light cannot be the white dwarf, the compact hotspot where the gas stream from the red dwarf feeds matter to the accretion disk, or any other compact region on the accretion disk. More detailed modelling of the eclipse observations will refine our knowledge of the spatial distribution of the superhump source, but this programme is complicated by the need to disentangle changes in the eclipse shapes from changes in the shape of the superhump. The slow march of the eclipses across the superhump light curve occurs on a time scale comparable with that for secular evolution of the superhump light curve itself.

In the starspot model proposed by Whitehurst *et al.*, superhumps are caused by a bright active region on the red dwarf. The discrepancy between the orbital and superhump periods is interpreted as a slight prograde rotation of the red dwarf in the frame of the orbital motion of the binary. The bright starspot must lie near the stellar equator to avoid occultation at $\phi_{\text{orb}} = 0.5$ by the accretion disk, because secondary eclipses are not observed.

Synthetic light curves for the starspot model are compared in Figs 3 and 4 with the observed superhump light curves of OY Car and Z Cha. Observations of OY Car by Kzreminski⁷ are shown in Fig. 3*a*. Warner's observations of Z Cha, as given by Cook⁸, are shown in Fig. 4*a* (also Fig. 2 of Whitehurst *et al.*). The light curves show typical examples of the large superhumps that develop a few days after the start of a superoutburst. A constant, equal to the depth of the eclipse, is subtracted from the original data to remove the light of the accretion disk. The amplitude of the brightness variation to be accounted for by the rotating

spotted star is thereby underestimated, as part of the light visible at mid-eclipse is, in fact, from the accretion disk. The accretion disk eclipses, which occurred at $\phi_{\text{sh}} \approx 0.5$ for Z Cha and $\phi_{\text{sh}} \approx 0.6$ for OY Car, are removed by interpolating observations before and after the eclipse.

Light curves for the starspot model are computed by summing contributions from an array of discrete sources distributed over the surface of the rotating star. The flux contributed by each visible source is given by $f = f_0 \cos \theta [1 - B + B \cos \theta]$, where θ is the angle between the line-of-sight and the normal to the stellar surface, f_0 is the flux when viewed normally, and B is the limb darkening parameter. The first $\cos \theta$ factor accounts for foreshortening, while the term in brackets approximates limb darkening for normal stellar atmospheres. The synthetic light curves shown in Figs 3*a* and 4*a* are for a ring of 100 fully darkened ($B = 1$) sources equally spaced around the equator of a spherical star.

The fluxes f_0 are constrained to be positive but otherwise are free to vary in fitting the observed light curve. Equatorial brightness distributions corresponding to the light curve fits in Figs 3*a* and 4*a* are given in Figs 3*b* and 4*b*. In the effort to produce a sharp peak corresponding to the superhump, the stellar brightness distribution breaks up into five precisely arranged bright regions. The structure in the synthetic light curves around $\phi_{\text{sh}} = 0.5$ is also an artefact of this failed attempt to produce a sharp peak. Calculations with sources covering the full stellar surface, with limb darkening parameters in the range $0 < B < 1$, using the correct Roche geometry, and including the eclipse of the red dwarf by the accretion disk, give similar results.

All of the starspot models considered here fail to produce light curves as sharply peaked as the observed superhumps. The starspot models fail because the light curve for a point source on the stellar equator, shown in Fig. 3*a* for undarkened ($B = 0$) and fully darkened ($B = 1$) cases, is already broader than the observed superhump. Any extended brightness distribution must produce an even broader peak. A bright spot at high stellar latitude produces a narrower peak by being visible for less than half of the rotation period, but such spots are ruled out by the absence of secondary eclipses.

The starspot model is thus incompatible with the observed superhump light curves, unless radiation from the starspot is beamed to a considerably greater extent than the radiation from a normal stellar atmosphere. The sharply peaked superhump light curves require the starspot radiation to emerge with an angular full width at half intensity of about 35° , compared with 90° – 120° for limb-darkened stellar atmospheres. As superhumps are seen in both eclipsing and non-eclipsing systems, this beaming must apply only to the direction of stellar longitude. Thus if superhumps arise on the red dwarf, some mechanism must be found to confine the radiation from the bright starspot to a narrow fan beam with its long axis in the direction of stellar latitude.

Light curve synthesis techniques have been used here to clarify the direct constraints that are placed by the light curves of eclipsing SU UMa systems on possible sites for superhumps. These constraints apply without regard to the physical mechanism responsible for superhumps, because they depend only on the well-understood geometry of eclipses in a binary system. Superhump sites at high stellar latitude on the red dwarf are eliminated by the absence of secondary eclipses at $\phi_{\text{orb}} = 0.5$. Sites near the stellar equator are feasible only by invoking a narrow fan beam radiation pattern to explain the sharply pointed shapes of superhumps together with the occurrence of superhumps in both eclipsing and non-eclipsing systems. The shapes of superhumps do not directly constrain disk models, in which the superhumps are produced by time variation rather than by occultation of the superhump source. The absence of deep eclipses of the superhump light in Z Cha, and the shapes of eclipses that occur during the superhump peak independently rule out compact sources on the accretion disk. There is no apparent conflict between the extant eclipse observations and a superhump source that is broadly distributed over the face of

the accretion disk. A more delicate analysis of the eclipse light curves should sharpen the observational constraints on the distribution of superhump light on the face of the accretion disk.

I thank G. Bath, B. Warner and J. Pringle for useful discussions.

Received 5 July; accepted 10 September 1984.

1. Vogt, N. *Astr. Astrophys.* **88**, 66–76 (1980).
2. Warner, B. Preprint (NATO Advanced Studies Institute, Cambridge, 1983).
3. Papaloizou, J. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **189**, 293–297 (1979).
4. Vogt, N. *Astrophys. J.* **252**, 653–667 (1982).
5. Whitehurst, R., Bath, G. T. & Charles, P. A. *Nature* **309**, 768–770 (1984).
6. Smak, J. *Acta astr.* (submitted).
7. Kzreminski, V. & Vogt, N. *Astrophys. J.* (submitted).
8. Cook, M. C. thesis, Univ. Cambridge (1982).

Differentiation between ^{27}Al nuclei in alums by Fourier transform pure nuclear quadrupole resonance

D. B. Zax, A. Bielecki & A. Pines

Department of Chemistry and Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720, USA

S. W. Sinton

Exxon Production Research Company, PO Box 2189, Houston, Texas 77001, USA

Aluminium is the most abundant metal in the Earth's lithosphere and is present in many commercially important materials. Recent studies on its role in zeolites¹ and proposed solar energy storage schemes involving aluminium salts² have led to renewed interest in analytical methods capable of identifying and characterizing specific crystalline sites. In particular, new techniques are necessary which do not require single crystals. One such technique is Fourier transform nuclear quadrupole resonance (NQR) with pulsed field cycling^{3,4}, and we report here on its application to the determination of specific ^{27}Al sites in model aluminium compounds, the alums.

In crystalline compounds, aluminium occurs most often with either four or six nearest neighbours. Tetrahedrally coordinated aluminium is the catalytically active centre in zeolites, where each ^{27}Al atom is believed to be surrounded by four O–Si linkages. The octahedral form, as characteristically found in the alum family, surrounds each ^{27}Al site by six bonds (often to H_2O molecules). Much work has gone into characterizing these sites. In powder samples, the NMR spectra are very broad and the methods used to obtain information have been two-dimensional solid state NMR⁵ and magic angle sample spinning (MASS) NMR⁶. Although the latter technique has proved useful in studies of ^{29}Si in zeolites^{7,8}, its applicability to aluminium is limited because of the small dispersion in ^{27}Al chemical shifts and the line broadening arising from quadrupolar effects, with the result that it generally reveals only the difference between tetrahedral or octahedral sites^{9,10}, except in rare cases of resolution between tetrahedral sites¹¹. As an example, Fig. 1 shows MASS spectra of ^{27}Al in two closely related materials, the ammonium and potassium alums ($\text{NH}_4\text{Al}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$ and $\text{KAl}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$). This technique reveals no difference between these two samples.

In contrast to high field NMR techniques, magnetic resonance should readily differentiate between signals from similar sites in a zero applied magnetic field; we rely in this instance on the inherent chemical sensitivity of the nuclear quadrupole interaction to resolve structural differences. The quadrupolar interaction arises from the coupling of a non-spherical nuclear charge distribution (characteristic of any nucleus where I , the nuclear angular momentum operator, is less than $\frac{3}{2}$) to molecular and crystalline electric field gradients. In ^{27}Al sites of high symmetry,

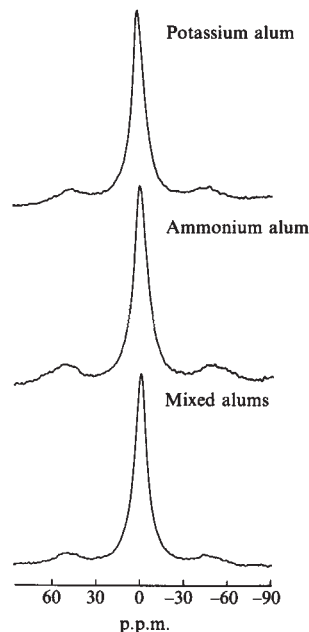


Fig. 1 ^{27}Al magic angle spinning spectra of representative alums. Spectra are observed at 78.2 MHz at a spinning rate of 4 kHz. Chemical shifts are referenced to $\text{Al}(\text{H}_2\text{O})_6^{3+}$. The mixed alum sample is 56 mol% potassium alum, 44 mol% ammonium alum.

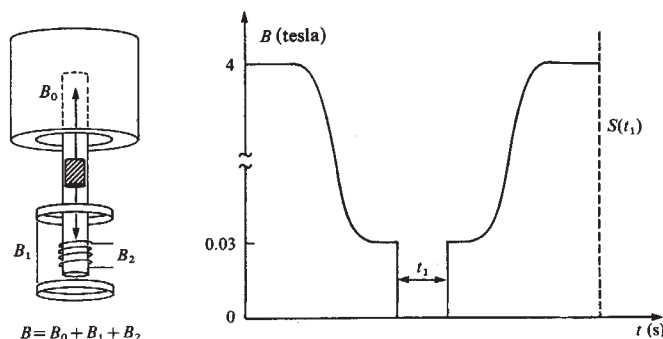


Fig. 2 Schematic diagram showing the effective field at the sample as a function of time. The sample is pneumatically transported from the bore of superconducting magnet B_0 to a point where B_1 accurately compensates for the fringe field. Coil B_2 (0.03 T) is switched off for a time t_1 during which pure quadrupolar evolution occurs. After returning the sample to B_0 , the high field magnetization, $S(t_1)$, is measured. The cycle is repeated for equally spaced points in t_1 until no further change in magnetization is observed. Fourier transformation of the magnetization as a function of t_1 results in the zero field spectrum.

the magnitude of the field gradient term is small. Direct detection of the pure NQR transitions is thus difficult (as the energy splittings are small¹²).

We instead used the chemical sensitivity of pure NQR, in combination with the high signal-to-noise ratio associated with observation of NMR in large applied magnetic fields. This is done by cycling the magnetic field between a large value and zero, and is accomplished in two steps (see Fig. 2): first, the sample is mechanically transported from the bore of a superconducting magnet and then the residual field is pulsed to zero with electronically switched coils. If the second stage occurs quickly relative to the quadrupolar frequencies, the spin system in zero field oscillates at frequencies characteristic of the pure quadrupolar coupling. This oscillation is interrupted at time t_1 by reapplying the field. The sample is then returned to high field where the signal amplitude is monitored. This field cycle is

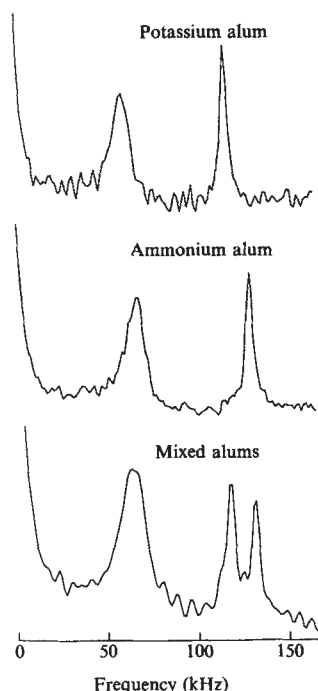


Fig. 3 ^{27}Al Fourier transform pure nuclear quadrupole resonance spectra of representative alums. Each site contributes two lines to the zero field spectrum. η , asymmetry parameter in the quadrupole tensor, was calculated to be 0.17 for potassium alum and ammonium alum. Respective values of 391 and 438 kHz were obtained for the quadrupole coupling constant, $(e^2qQ)/h$, where eQ is the nuclear quadrupole moment, eq is the electric field gradient, and h is Planck's constant. The pair of high frequency lines in the mixed sample (56 mol% potassium alum, 44 mol% ammonium alum) clearly indicate the existence of two distinct ^{27}Al sites.

repeated with different, equally spaced values of t_1 , to give an interferogram which describes the evolution of the spin system in zero field and which, on Fourier transformation, yields the zero field spectrum.

For a spin- $\frac{5}{2}$ nucleus such as ^{27}Al , we expect to observe a pair of lines for each crystalline site. (In aluminium compounds where the crystal symmetry is low, a third line of low intensity may be observed.) Figure 3 illustrates the results for the two alums whose MASS spectra are shown in Fig. 1. Despite nearly identical crystal structures and high field NMR spectra, the zero field nuclear quadrupole experiment clearly differentiates between the two types of aluminium nuclei. We expect that even more sensitive differentiation will be possible between ^{27}Al sites in other systems such as zeolites.

D.B.Z. gratefully acknowledges support by a NSF predoctoral fellowship. This project was supported by the Center for Advanced Materials and the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Science Division of the US Department of Energy under contract DE-AC03-76SF00098.

Received 1 August; accepted 5 September 1984.

1. Barrer, R. M. *Hydrothermal Chemistry of Zeolites*, Ch. 1 (Academic, London, 1982).
2. Vaccarino, C. & Fioravanti, T. *Sol. Energy* **30**, 123-125 (1983).
3. Bielecki, A. *et al. J. chem. Phys.* **80**, 2232-2234 (1984).
4. *Chemical and Engineering News* **61** (50), 23-24 (1983).
5. Samoson, A. & Lippmaa, E. *Chem. Phys. Lett.* **100**, 205-208 (1983).
6. Meadows, M. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 1351-1355 (1982).
7. Lippmaa, E., Mägi, M., Samoson, A., Tarmak, M. & Engelhardt, G. *J. Am. chem. Soc.* **103**, 4992-4996 (1981).
8. Thomas, J. M., Klinowski, J., Ramdas, S., Hunter, B. K. & Tennakoon, D. T. *B. Chem. Phys. Lett.* **102**, 158-162 (1983).
9. Müller, D., Gessner, W., Behrens, H.-J. & Scheler, G. *Chem. Phys. Lett.* **79**, 59-62 (1981).
10. Fyfe, C. A., Gobbi, G. C., Klinowski, J., Thomas, J. M. & Ramdas, S. *Nature* **296**, 530-533 (1982).
11. Klinowski, J., Anderson, M. W. & Thomas, J. M. *JCS chem. Commun.* 525-526 (1983).
12. Das, T. P. & Hahn, E. L. *Solid St. Phys. Suppl.* **1**, 70-77 (1958).

Direct determination of strontium enrichment on grain boundaries in a garnet lherzolite xenolith by proton microprobe analysis

Donald G. Fraser*, F. Watt†, G. W. Grime† & J. Takacs†

* Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK

† Department of Nuclear Physics, University of Oxford, Keble Road, Oxford, OX1 3RH, UK

The mechanisms by which trace and radiogenic elements are held and transported in the Earth's mantle are of prime importance in studies of basalt petrogenesis and the chemical processing of the mantle. While such mechanisms include the migration of melts and the convective physical admixture of differentiated material¹, the role of fluids as transport agents has attracted much interest, especially, in regions affected by kimberlitic activity²⁻⁴, continental basaltic volcanism^{5,6}, and in areas of subduction-related magmatic activity^{7,8}. The observation that many incompatible elements are readily leached from garnet lherzolites by dilute acid treatment³ has led to the assumption that a proportion of these elements is located on grain boundaries, rather than in solution in the various minerals present. This indirect evidence is reinforced by comparison of the compositions of mineral separates with their bulk parental material^{4-8,11}. Although careful electron microprobe analyses have made it possible to detect minor and trace elements within grains¹², the low concentrations involved, together with the high X-ray background of the electron microprobe, have made it impossible to investigate the supposed distribution of incompatible trace elements on grain boundaries. We have measured directly, using a high-resolution proton microprobe^{9,10}, the distribution of Sr and other elements in a garnet lherzolite xenolith. The analyses demonstrate that, in this sample, Sr is preferentially concentrated along grain boundaries.

An inherent limitation in the use of the electron microprobe to analyse very, small concentrations of elements lies in the substantial non-characteristic X-ray background caused by electron *bremsstrahlung*. This is very much reduced when a proton beam is used to excite X rays in a sample because the enormously increased mass of the proton minimizes the amount of deceleration near the target. In the present experiments, the specimens were analysed by using a 4.0-MeV proton beam to excite characteristic X rays in the samples. The beam was focused electromagnetically to a spot diameter of 2 μm and was rastered over an area of up to 4 mm $\times$ 4 mm. The samples were prepared as 30- μm thin sections and the X rays generated were detected and analysed using a Link Si(Li) detector system. As the penetration of a proton beam is much greater than that of a beam of electrons in an electron microprobe, the sections were mounted on glass slides using Lakeside 70 thermoplastic resin as a mounting medium. The sections were then removed from the glass slides by gentle heating and remounted over 1-cm holes machined in aluminium target holders to remove any possibility of X-ray generation in the specimen backing. The resulting targets were finally washed carefully in acetone to remove any traces of resin. Samples were irradiated for ~ 10 h using a 50 μm Al filter to cut out the low-energy X-ray background, and maps of the distribution of X rays from major and trace elements in the samples were built up as 256 $\times$ 256 pixel images. As the proton beam penetrated right through the thin samples used, the transmitted beam current was also continuously measured for each pixel so that the effects of any fluctuations in beam current could be corrected. A detailed description of the Oxford proton microprobe system is given elsewhere^{9,10}.

Figure 1 shows elemental maps for the distribution of Ni, Mn, Cr and Sr in a 4 mm $\times$ 4 mm area of a granular garnet

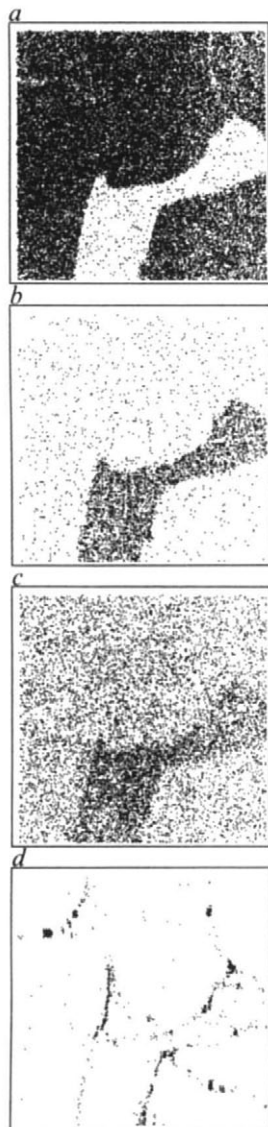


Fig. 1 Elemental X-rays maps of the distribution of Ni (a), Cr (b), Mn (c) and Sr (d) in a 4×4 mm thin section of a coarse garnet lherzolite xenolith obtained using the proton microprobe. Ni, Cr and Mn are present within the mineral grains, but Sr is concentrated on the grain boundaries. Dark areas indicate high concentrations of the elements.

lherzolite xenolith (MK 15) collected from the Bultfontein dumps. The images were displayed in false colour according to the number of X-ray counts per pixel, but are reproduced here as black and white photographs in which the dark areas represent the highest concentrations. The section consists of an angled grain of orthopyroxene (Cr- and slightly Mn-rich and Ni-poor) surrounded by Ni-rich olivine, and there is no obvious zoning in the concentrations of these elements. In marked contrast to the distribution of Cr, Mn and Ni, the map of Sr distribution shows almost no Sr in the interiors of the grains; instead, Sr is concentrated along grain boundaries in the specimen. The Sr was clearly introduced into the rock after its present texture had formed and comparison with the thin section shows that it is located on the serpentinized grain boundaries and cracks in the mineral grains.

To provide an estimate of the concentration of Sr represented in Fig. 1, the data were calibrated by analysing samples of clinopyroxenes which had been analysed previously as mineral separates by isotope dilution¹³. The standards were also prepared as 30-μm thin sections and data were collected from each standard as the number of Sr-K_α counts per nanocoulomb of beam charge deposited on the target. A plot of the count rates against the Sr-contents of the three clinopyroxenes used as standards is given in Fig. 2. The Sr concentrations along the grain boundaries of the garnet lherzolite xenolith, shown in Fig. 1, produced 6–36 counts per nanocoulomb per pixel corresponding therefore to a range of 100–600 p.p.m.

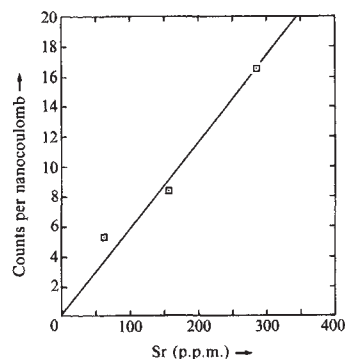


Fig. 2 Calibration of Sr count rates. The number of counts per nanocoulomb of deposited proton charge is plotted against the Sr concentrations of clinopyroxenes which had been analysed for Sr by isotope dilution.

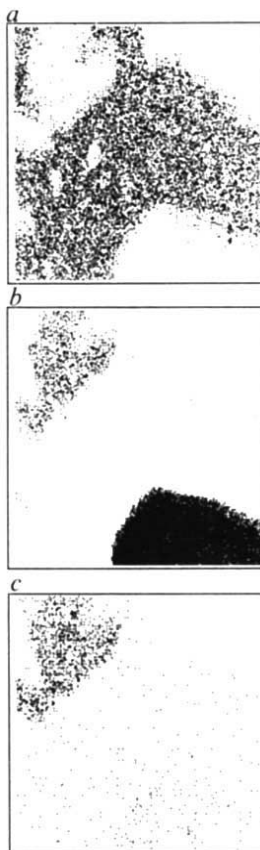


Fig. 3 Elemental X-ray maps of the distribution of Ni (a), Cr (b) and Sr (c) in a thin section of a mosaic porphyroclastic garnet lherzolite xenolith. In this specimen, Sr is distributed like the other trace elements within mineral grains and there is no sign of the process which caused the concentration of Sr on grain boundaries in the specimen in Fig. 1.

Although it is highly unlikely that Sr at these levels could have been introduced into the sample during preparation of a thin section, a second garnet lherzolite sample was selected and prepared in the same way for analysis. This sample (BD 2501) had a mosaic porphyroclastic texture¹⁴, and was known, on the basis of previous studies¹³, to be, of all the xenoliths, one of the least affected by acid leaching. Elemental maps were built up as before and the images for Sr, Ni and Cr in a 4 mm×4 mm section are shown in Fig. 3. The area analysed contains clinopyroxene and garnet grains at the top and bottom of the photographs respectively, separated by a larger Ni-rich olivine grain. The results are in marked contrast to the data shown in Fig. 1 and indicate that, in this sample, there is no detectable concentration of Sr on grain boundaries. Instead, the Sr is present within the mineral grains, the majority being in the clinopyroxene, but with detectable levels in the garnet also. The absence of Sr located along grain boundaries in this section, therefore, makes it unlikely that the data shown in Fig. 1 are an artefact of sample preparation.

These preliminary results provide direct evidence for the location of Sr on serpentinized grain boundaries in certain samples from the upper mantle and strongly suggest the role of an aqueous fluid as the transporting agent. They also indicate the success of the high-resolution proton microprobe as an instrument for determining the concentrations and spatial distribution of trace elements in geological samples. Detailed isotopic work on the timing and nature of the event responsible for the introduction of Sr into the xenolith shown in Fig. 1 will be presented elsewhere (D.G.F. and S. Beswetherick, in preparation).

D.G.F. acknowledges financial support from the NERC and the Anglo-American Corporation. We thank Dr S. Beswetherick for providing the analysed specimens used as calibration points.

Received 18 June; accepted 14 September 1984.

1. McKenzie, D. & O'Nions, R. K. *Nature* **301**, 229-231 (1983).
2. Erlank, A. J. & Shimizu, N. *Abstr. 2nd int. Kimberlitic Conf.*, Santa Fe (1977).
3. Erlank, A. J., Allsopp, H. L., Hawkesworth, C. J. & Menzies, M. A. *Terra Cognita* **2**, 261-263 (1982).
4. Carswell, D. A. *Phys. Chem. Earth* **9**, 417-429 (1975).
5. Cox, K. G. in *Continental Basalts and Mantle Xenoliths* (eds Hawkesworth, C. J. & Norry, M. J.) 139-157 (Shiva, Cheshire, 1983).
6. Hawkesworth, C. J., Erlank, A. J., Marsh, J. S., Menzies, M. A. & van Calsteren, P. in *Continental Basalts and Mantle Xenoliths* (eds Hawkesworth, C. J. & Norry, M. J.) 111-138 (Shiva, Cheshire, 1983).
7. Stern, C. R. & Wyllie, P. J. *Am. Miner.* **63**, 641-663 (1978).
8. Best, M. G. *Geology* **3**, 429-432 (1975).
9. Grime, G. W. & Watt, F. *Beam Optics of Quadrupole Probe-forming Systems*, 1-263 (Hilger, Bristol, 1984).
10. Shimizu, N. *Phys. Chem. Earth* **9**, 655-669 (1975).
11. Erlank, A. J. *Yb. Carnegie Instn Wash.* **68**, 433-438 (1970).
12. Beswetherick, S. thesis, Univ. Oxford (1984).
13. Harte, B. *Geology* **85**, 279-288 (1977).

Upwarp of anomalous asthenosphere beneath the Rio Grande rift

E. C. Parker*, P. M. Davis*, J. R. Evans†, H. M. Iyer† & K. H. Olsen‡

* Department of Earth and Space Sciences, University of California, Los Angeles, Los Angeles, California 90024, USA

† US Geological Survey, Menlo Park, California 94025, USA

‡ Institute of Geophysics and Planetary Physics, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

Continental rifts are possible analogues of mid-ocean ridges, although major plate tectonic features are less clearly observed¹. Current thermal models of mid-ocean ridges²⁻⁴ consist of solid lithospheric plates overlying the hotter, less viscous asthenosphere, with plate thickness increasing away from the ridge axis. The lithospheric lower boundary lies at or near the melting point isotherm, so that at greater depths higher temperatures account for lower viscosity, lower seismic velocities and possibly partial melting. Upwarp of this boundary at the ridge axis concentrates heat there, thus lowering densities by expansion and raising the sea floor to the level of thermal isostatic equilibrium. At slow spreading ridges, a major central graben forms owing to the mechanics of magma injection into the crust⁵. Topography, heat flow, gravity and seismic studies support these models. On the continents, a low-velocity channel has been observed, although it is poorly developed beneath ancient cratons⁶⁻⁹. Plate tectonic models have been applied to continental basins and margins¹⁰⁻¹², but further similarities to the oceanic models remain elusive. Topographic uplift is often ascribed to Airy type isostatic compensation caused by crustal thickening, rather than thermal compensation in the asthenosphere. Here we discuss the Rio Grande rift, in southwestern United States. Teleseismic P-wave residuals show that regional uplift is explained by asthenosphere uplift rather than crustal thickening.

The Rio Grande rift extends for at least 950 km from central Colorado through New Mexico into Mexico. It separates the Great Plains to the east from the Colorado plateau and basin

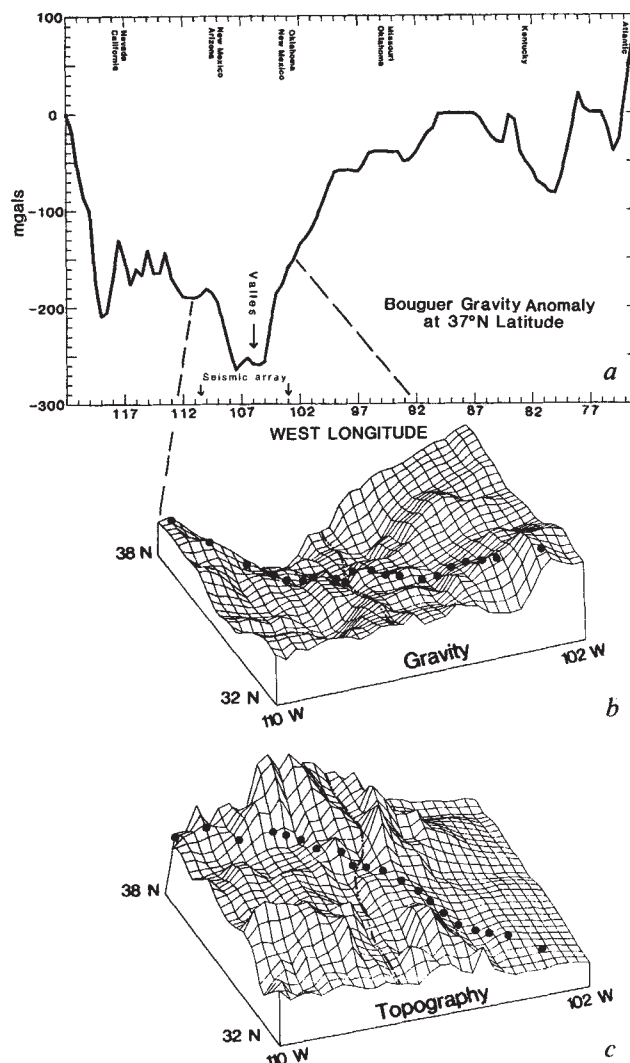


Fig. 1 *a*, East-west Bouguer gravity profile across the United States at latitude 37° N. The Valles caldera and Rio Grande rift lie within the most negative trough. *b*, Three-dimensional perspective view of Bouguer gravity showing locations of teleseismic stations, and Rio Grande River (---). *c*, Three-dimensional perspective of topography showing regional uplift, which is inversely correlated to the gravity of *b*.

and range province to the west and lies just east of the continental divide¹³. Figure 1 shows the transcontinental Bouguer anomaly at 37° N latitude, which crosses the rift along the New Mexico-Colorado border. The rift lies in a region containing the most negative Bouguer anomalies in the continental United States. Figure 1*b* and *c* are three-dimensional perspectives of the regional gravity and topography, which at long wavelengths are inversely correlated. Our experiment investigates whether the basis of the correlation is crustal or asthenospheric.

Rayleigh wave depths to the continental low-velocity zone of the asthenosphere range from about 70 km (refs 8, 9) in the Great Basin of Nevada and western Utah to 83-145 km (ref. 7) in the south-central part of the continent. Although the gravity and topography variations of Fig. 1*b, c* have been attributed^{14,15} to such undulation of lithospheric thickness, the boundary shape has not been determined directly. To probe the Earth to this depth, a high-energy seismic source is required. This is most conveniently supplied by distant earthquakes in the form of teleseismic waves.

The teleseismic P-wave technique¹⁶⁻¹⁹ consists of timing arrivals from distant earthquakes and subtracting from them tabulated²⁰ arrival times for a laterally homogeneous Earth. Any residuals remaining are attributed to lateral heterogeneity

beneath the seismic array, because distant heterogeneities, such as those at the source, will be attenuated by geometric spreading. Incidence and azimuthal angles are calculated from source parameters using a standard Earth model. The pattern of residuals has been observed to shift in response to different incidence and azimuthal angles in the incoming waves. Given enough rays at different angles, the data can be inverted to determine velocity at depths of up to about half the array dimension, with a resolution equal to station spacing near the surface but decreasing with depth.

Figure 2 shows the locations of the 20 recording stations along the 1,000-km line used for the 1982–83 experiment. A north-west–south-east orientation was taken because teleseismic arrivals lie predominantly along this azimuth so that rays from the north-west and south-east intersect in regions lying beneath the array. The array ran from late November 1982 until late January 1983.

Of the 60 earthquakes recorded, 40 have sufficient signal-to-noise ratio to be used in an inversion. Twenty-five of these lie along the array azimuth. The relative P-wave arrival times can be measured to 0.1 s. Residuals are found by subtracting travel times, obtained from the Herrin²⁰ tables, from the arrival times. Relative residuals, formed for each event by subtracting the mean from the raw residuals, are plotted in Fig. 3a. Some of the scatter is due to changes of the residual pattern with the incidence and azimuthal angle.

Additional assumptions must be made if we are to find an unequivocal interpretation for the residual pattern. Therefore, we assume that the regional shifting pattern is mainly due to upwarp of the asthenosphere–lithosphere boundary, that localized heterogeneity gives rise to the small-scale variations, and that the boundary in the vicinity of our array is approximately two-dimensional with a north–south strike. We used synthetic models to test for a way of inverting for such a boundary and decided on two methods: (1) downward projection of the data to give the boundary shape directly²¹ and (2) the two-dimensional^{19,22,23} damped, least-squares block method to give the averaged velocity distribution. Both methods gave similar results (Fig. 3b, c).

The projection method assumes that the interface is not too deep and that velocity contrasts are not too extreme, so that a straight ray and plane wave approximation for the teleseismic waves is adequate. The arrival time residuals will then be proportional to the length of the ray in the anomalous velocity zone, with the proportionality dependent on the velocity contrast. The raw residual pattern gives a distorted picture of the interface—it is projected onto a rotated axis—but as we know the angle of incidence, we can transform the projection to regain the original shape. The transformed patterns are then projected down each ray and the position of maximum overlap determines the true depth. Maximum overlap was determined by minimizing the sum of squares of residuals to a sixth-order polynomial fit by varying both depth and velocity contrast. Figure 3b shows the best-fit model which has a depth range of 70–200 km and a P-wave velocity contrast of -8% relative to a value of 8 km s^{-1} above the boundary. The average misfit of the polynomial to the data was 0.3 s.

The block model involves dividing the region into two-dimensional blocks and determining the velocity perturbations that give the observed delays. After running a series of tests to determine the uniqueness and minimum parameterization of the block model and having found an optimal model, we tested the hypothesis that all the residuals are due to slowness located in the crust or upper lithosphere rather than at depth. The hypothesis was rejected at the 99.9% confidence level: the data require a deep low-velocity zone. Thus, the model we present (Fig. 3c) has interfaces at 40, 70 and 200 km. A block width of 100 km was chosen because our station spacing ranged from 30 to 140 km. The main feature of this model is a spatially coherent 500-km broad, low-velocity zone of -0.35% to -5.1% contrast, in the depth range 70–200 km. Smaller variations in the upper two layers are incoherent. Incoherency was also found

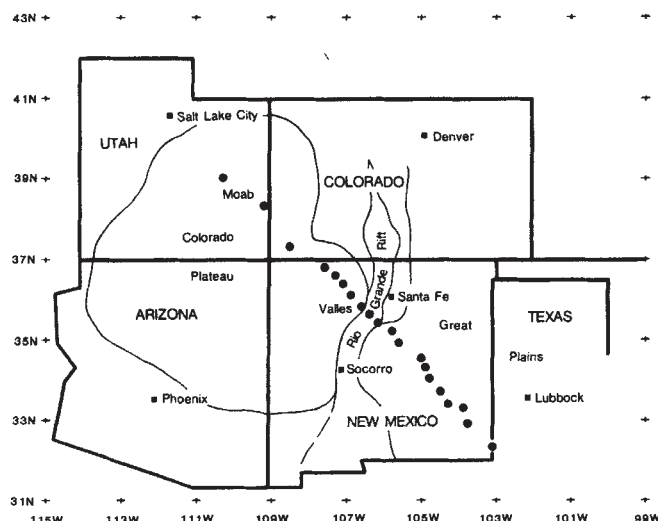


Fig. 2 Station locations. Each consists of a 1-Hz, L4C vertical seismometer and recorder.

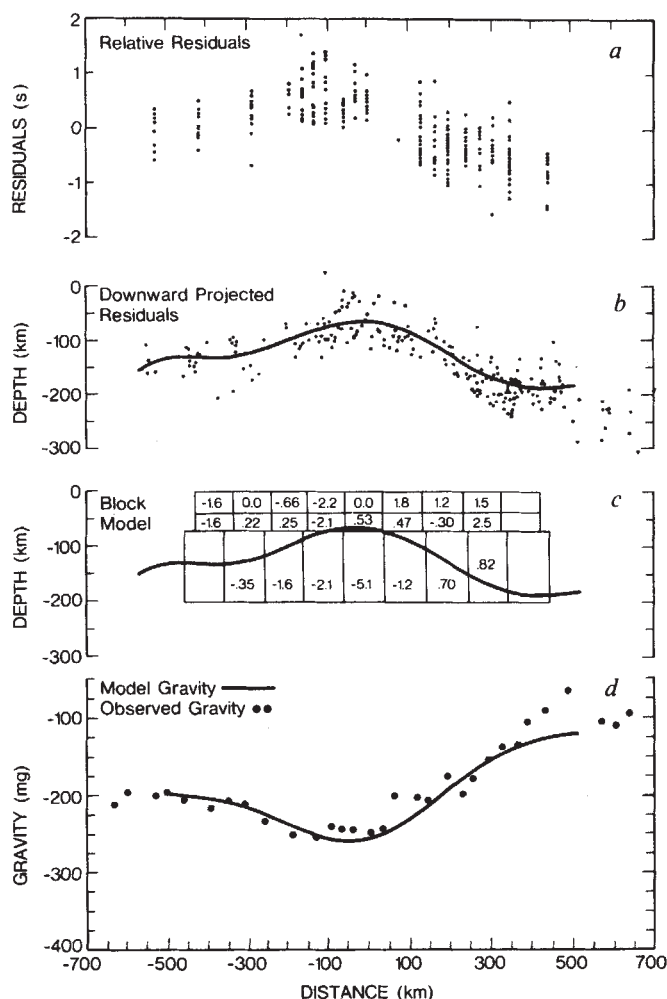


Fig. 3 a, Teleseismic P-wave residuals (de-measured) for 22 events. b, Downward projection of the residual patterns (of a) along ray paths, after rotation to correct for incidence effects. An approximate shape of the inferred lithosphere–asthenosphere boundary is given by the sixth-order polynomial fit. c, Block and polynomial model comparison. Numbers refer to per cent velocity changes compared with -8% contrast found for the polynomial model. Layer interfaces were taken at 40, 70 and 200 km, with unperturbed layer velocities of 6.5, 8.0 and 8.0 km s^{-1} , respectively. d, Gravity anomaly measured along the teleseismic array compared with gravity calculated from the polynomial model of b.

in an additional layer, 130 km thick, when added to the base of the model. The negative anomaly remained unaffected. Standard errors given by the least-squares solution are all $\sim 0.6\%$. The diagonal elements of the resolution matrix ranged from 0.55 in the top layer to 0.86 in the lowest layer. Average misfit for the final model is 0.297 s.

The Bouguer gravity anomaly calculated for both the polynomial model and the block model shows good agreement with the measured data for the area. Figure 3d shows the fit for the gravity values calculated from the polynomial model for which the adjusted density contrast in the low-velocity zone gave a value of -0.05 g cm^{-3} , that is, $\sim 1.6\%$ of 3.2 g cm^{-3} . The ratio of velocity to density contrast, 5.0, is much larger than expected from the Nafe-Drake²⁴ relationship determined from different types of igneous rock, which gives a value of ~ 0.8 . Such an increase in the sensitivity of velocity to change in density is consistent with the presence of a small fraction of partial melt within the low-velocity zone which, while having little effect on density, will have a disproportionately larger effect on shear modulus and hence P-wave velocity. High attenuation supports this view. P-wave amplitudes measured at Albuquerque, which lies on the rift, are among the most attenuated of those measured at standard stations in the United States²⁵. Greater attenuation is seen only at Golden, Colorado, to the north, on the Rocky Mountain front.

The topography also agrees qualitatively with that in isostatic balance with the polynomial model, that is, in thermal isostatic equilibrium. Therefore, the model we assume to explain the teleseismic delays, attenuation²⁵, gravity, uplift^{4,15}, regional heat flow²⁶ and the presence of rifting and volcanism, is based on a hot, low-density asthenospheric body which has replaced the lithosphere by upward movement of the lithosphere-asthenosphere boundary. Such migration of the melting point geotherm might be caused by either lithospheric extension¹⁰ or active convection²⁷ in the upper mantle and requires further quantitative work to elucidate the mechanism. Nonetheless, the boundary has a configuration expected to exist in a region as a prelude to the opening of oceans and continental drift. Although the polynomial fit of Fig. 3b shows some agreement with lithospheric depths determined from Rayleigh waves, other points are still unresolved. Where low P-wave velocities have been associated with low S-wave velocities^{9,28,29} in the low-velocity zone, the per cent reduction in P-velocity is about half that in S-velocity. Our model reduction in P-velocity is $\sim 8\%$, implying an unprecedented regional reduction of S-velocity of $\sim 14\%$. Also, the low-velocity zone has been associated with increased density values^{7,28}, which is opposite to our conclusion. If we compromise between less P-wave velocity perturbation and increased amplitude of the variation of the boundary, the flanks would be placed too deep for them to interface with surface wave models of lithospheric thickness. Therefore, if asthenospheric upwarp is to explain the gravity, topography and teleseismic delays, the properties of the asthenosphere in this region must differ from those encountered elsewhere. However, we might expect this, as it underlies the most negative region of the east-west, continental Bouguer gravity profile (latitude 37° , Fig. 1).

This work was supported by grants from the University of California, Los Angeles, and the Institute of Geophysics and Planetary Physics at Los Alamos National Laboratory.

Received 7 August; accepted 5 September 1984.

1. Morgan, P. & Baker, B. H. *Tectonophysics* **94**, 1-4 (1983).
2. Parker, R. L. & Oldenburg, D. W. *Nature* **242**, 137-139 (1973).
3. Oldenburg, D. W. *Geophys. J. R. astr. Soc.* **43**, 425-451 (1975).
4. Schubert, G., Froidevaux, C. & Yuen, D. A. *J. geophys. Res.* **81**, 3525-3540 (1976).
5. Davis, P. M. *Nature* **308**, 53-55 (1984).
6. Knopoff, L. *Tectonophysics* **13**, 497-519 (1972).
7. Biswas, N. N. & Knopoff, L. *Geophys. J. R. astr. Soc.* **36**, 515-540 (1974).
8. Priestley, K. & Brune, J. J. *geophys. Res.* **83**, 2265-2272 (1978).
9. Priestley, K., Orcutt, J. A. & Brune, J. N. *J. geophys. Res.* **85**, 7166-7174 (1980).
10. McKenzie, D. *Earth planet. Sci. Lett.* **40**, 25-32 (1978).
11. Sleep, N. H. *Geophys. J. R. astr. Soc.* **24**, 325-350 (1971).
12. Steckler, M. S. & Watts, A. B. in *Dynamics of Passive Margins, Geodynamics Ser. Vol. 6*, 184-196 (Am. Geophys. Un., Washington, D.C., 1982).
13. Ramberg, I. B., Cook, F. A. & Smithson, S. B. *Bull. geol. Soc. Am.* **89**, 107-123 (1978).

14. Cordell, L. *Bull. geol. Soc. Am.* **89**, 1073-1090 (1978).
15. Ander, M. E. thesis, Univ. New Mexico (1980).
16. Iyer, H. M. *Nature* **253**, 425-427 (1975).
17. Aki, K. *Rev. Geophys. Space Phys.* **20**, 161-170 (1982).
18. Iyer, H. M. *Phil. Trans. R. Soc. A310*, 473-510 (1984).
19. Evans, J. R. *J. geophys. Res.* **87**, 2654-2670 (1982).
20. Herrin, E. *Bull. seism. Soc. Am.* **58**, 1196-1219 (1968).
21. Davis, P. M., Parker, E. C., Evans, J. R., Iyer, H. M. & Olsen, K. H. in *New Mexico Geological Society Guidebook* (New Mexico Geol. Soc., Socorro, New Mexico, 1984).
22. Aki, K., Christofferson, A. & Husebye, E. S. *Bull. seism. Soc. Am.* **66**, 501-524 (1976).
23. Ellsworth, W. A. thesis, Massachusetts Inst. Technol. (1977).
24. Ludwig, W. J., Nafe, J. E. & Drake, C. L. *The Sea Vol. 4*, 53-84 (Wiley, New York, 1970).
25. Butler, R. *Rev. Geophys. Space Phys.* **22**, 1-36 (1984).
26. Rieter, M., Mansure, A. J. & Shearer, C. in *Rio Grande Rift: Tectonics and Magmatism* (ed. Rieker, R. E.) 253-268 (Am. Geophys. Un., Washington, D.C., 1979).
27. Sengor, A. M. C. & Burke, K. *Geophys. Res. Lett.* **5**, 419-421 (1978).
28. Dziewonski, A. M., Hales, A. L. & Lapwood, E. R. *Phys. Earth planet. Inter.* **10**, 12-48 (1975).
29. Burdick, L. J. & Helmberger, D. V. *J. geophys. Res.* **83**, 1699-1712 (1978).

Enhanced preservation of marine-derived organic matter in Cenomanian black shales from the southern Angola Basin

P. A. Meyers*, M. J. Leenheer†, O. E. Kawka*‡ & T. W. Trull*‡

* Department of Atmospheric and Oceanic Science, University of Michigan, Ann Arbor, Michigan 48109, USA

† Cities Service Technology Center, PO Box 3908, Tulsa Oklahoma 74102, USA

Black shales possessing high concentrations of organic carbon¹ were deposited in many parts of the proto South Atlantic Ocean during the Cretaceous period². The way such sediments accumulated is not fully understood, but is likely to have occurred through a combination of low oxygen availability and abundant supply of organic matter. Thin, centimetre-thick layers of black shales are commonly interbedded with thicker layers of organic carbon-deficient, green claystones, as found in strata of Aptian to Coniacian age, at Deep Sea Drilling Project (DSDP) Site 530, in the southern Angola Basin³ and elsewhere. These differences in carbon content and colour reflect the conditions of deposition, and possibly variations in the supply of organic matter^{4,5}. We have compared, using organic geochemical methods the compositions of organic matter in three pairs of closely-bedded black and green Cenomanian claystones obtained from Site 530. Kerogen analyses and distributions of biological markers show that the organic matter of the black shales is more marine and better preserved than that of the green claystones.

The results of the analyses are presented in Table 1. X-ray diffraction studies do not show any significant differences between the mineralogical contents of the black and green claystones. Furthermore, the pyrite content of adjacent intervals is similar: its relative abundance seems to depend more on the depositional setting than on the amount of organic matter in the sediments.

Values of total organic carbon, however are significantly higher for the black shales than for their green claystone pairs. Modern deep-sea sediments generally contain $\sim 0.2\%$ organic carbon⁶. Values as high as in these black shales are found in only a few areas of present-day oceans⁷, which are under anoxic or poorly oxygenated bottom-waters. A rare combination of abundant supply and exceptionally good preservation of organic matter is needed for such organic carbon-rich sediments to accumulate.

Visual kerogen analysis of the green and black claystones shows that both contain two populations of vitrinite and amorphous kerogen. Whereas the indigenous vitrinite content indicates that the samples are thermally immature, Fig. 1 shows that the green claystones differ from the black shales in containing a larger amount of recycled vitrinite having a higher thermal

‡ Present addresses: School of Oceanography, Oregon State University Corvallis, Oregon 97301, USA (O.E.K.); Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA (T.W.T.).

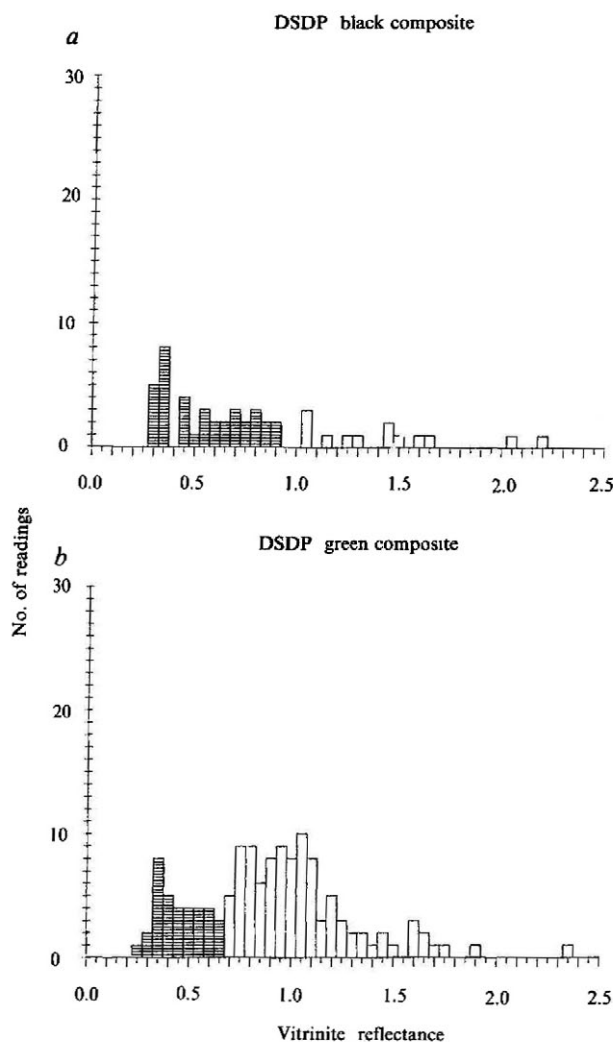


Fig. 1 Vitritine reflectance histograms of composite samples of Cenomanian black shale (*a*, mean = 0.545, median = 0.535, s.d. = 0.195, count = 37) and green claystone (*b*, mean = 0.457, median = 0.450, s.d. = 0.113, count = 35) from DSDP Site 530 in the Angola Basin. Hatched area represents indigenous population.

maturity. Similarly, populations of spores present in the green claystones range from transparent to dark orange, indicating a wide range of thermal maturities, whereas the black claystones contain primarily light-coloured, thermally immature spores. Also the amorphous kerogen of the green claystone is ~85% fluorescent compared with 100% in the black shales. Because thermally immature amorphous kerogen is highly fluorescent, the lower percentages in the green claystones suggest they contain a significant amount of nonfluorescing detrital material. The presence of two populations of both vitritine and amorphous kerogen, indicated as in Fig. 1 by the vitritine reflectance histograms, and by visual inspection, implies there are differences in the organic matter intake and preservation of the two claystones.

The mean value of organic matter C/N atomic ratios is much greater for the black shales (38.1) than for the green claystones (5.7) and is typical of C/N ratios of cellulosic land-plant organic matter⁸. The ¹³C content of organic matter in black shales from Site 530 is generally less than in green claystones. This difference in isotopic compositions could mean there is a larger component of land-derived organic material in the black than in the green strata. Characterization of organic matter by kerogen analysis, however, shows that the black shales of the Angola Basin contain more amorphous kerogen, indicative of their aquatic origin, than do the green claystones¹⁰⁻¹².

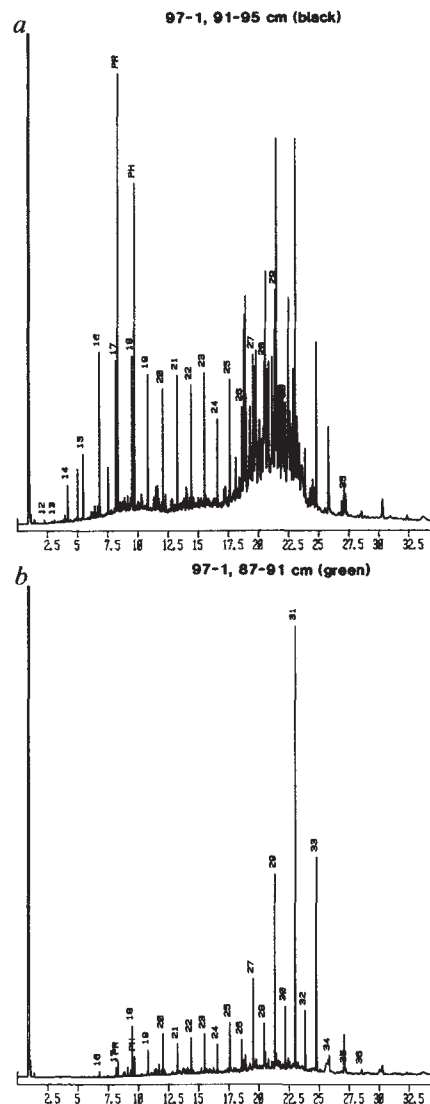


Fig. 2 Gas chromatograms showing extractable aliphatic hydrocarbon distributions of adjacent samples of black (*a*) and green (*b*) claystones from Hole 530A. Pristane and phytane are identified as PR and PH, respectively, and *n*-alkanes by their carbon numbers.

The contradictions in the identification of the source of organic matter, apparent between the C/N ratios and carbon isotope content and the kerogen analyses may arise from selective preservation of organic matter components. Although preservation is generally high in the black shales, nitrogenous materials appear to have been selectively lost, thus enhancing C/N ratios. Similarly, classes of organic compounds relatively rich in ¹³C, such as amino acids and carbohydrates, may have been selectively degraded, as reported in sediments from Mangrove Lake³, resulting in reduced ¹³C in the residual organic matter of the black shales. Poorer preservation in the green claystones has led to losses of all but the most resistant forms of organic matter.

Geolipid data agree with this picture of selective preservation. Concentrations of total hydrocarbons, fatty acids, and fatty alcohols are significantly greater in the black shales than in the adjacent claystone (Table 1), yet these geolipids often constitute a smaller fraction of the total organic matter in black shales than in green claystones¹⁴, which is consistent with greater preservation of the more degradable, non-lipid organic matter components in the black shales.

The hydrocarbons in the samples are a combination of compounds either introduced as such to the sediments or generated during diagenesis. The large amounts of aliphatic and aromatic hydrocarbons, in contrast to the smaller amounts of alkanols

Table 1 Comparisons of adjacent Cenomanian black shales with green claystones from DSDP Hole 530A in the Angola Basin

Characteristic of sample	Sample pairs					
	530A-95-5		530A-97-1		530-97-1	
	0-21 cm	38-50 cm	87-91 cm	91-95 cm	99-105 cm	105-110 cm
Colour	Green	Black	Green	Black	Black	Green
% Calcium carbonate	2	9	5	7	6	6
% Pyrite	4	17	6	6	14	16
% C _{organic}	0.2	12.9	0.6	4.9	10.0	0.8
Atomic C/N	4.5	36.0	6.8	38.1	40.3	5.7
% Amorphous kerogen	60	98	60	98	95	90
% Vitrinite-intertinite	40	2	40	2	5	10
Aliphatic hydrocarbons (p.p.m.)	4	274	20	64	135	18
Aromatic hydrocarbons (p.p.m.)	6	695	20	202	284	126
Alkanoic acids (p.p.m.)	3	7	2	6	11	3
Alkanols (p.p.m.)	0.4	5	2	5	9	3

Samples for geochemical comparison, taken from ~1,020 m sub-bottom depth, were selected onboard ship and immediately frozen to protect their organic matter. Analysis of organic carbon and of solvent-extractable geolipids were done as described elsewhere¹⁴. Kerogen was isolated by demineralization and flotation, and analysed microscopically in both transmitted and reflected light. Bulk mineralogy was determined by X-ray diffraction.

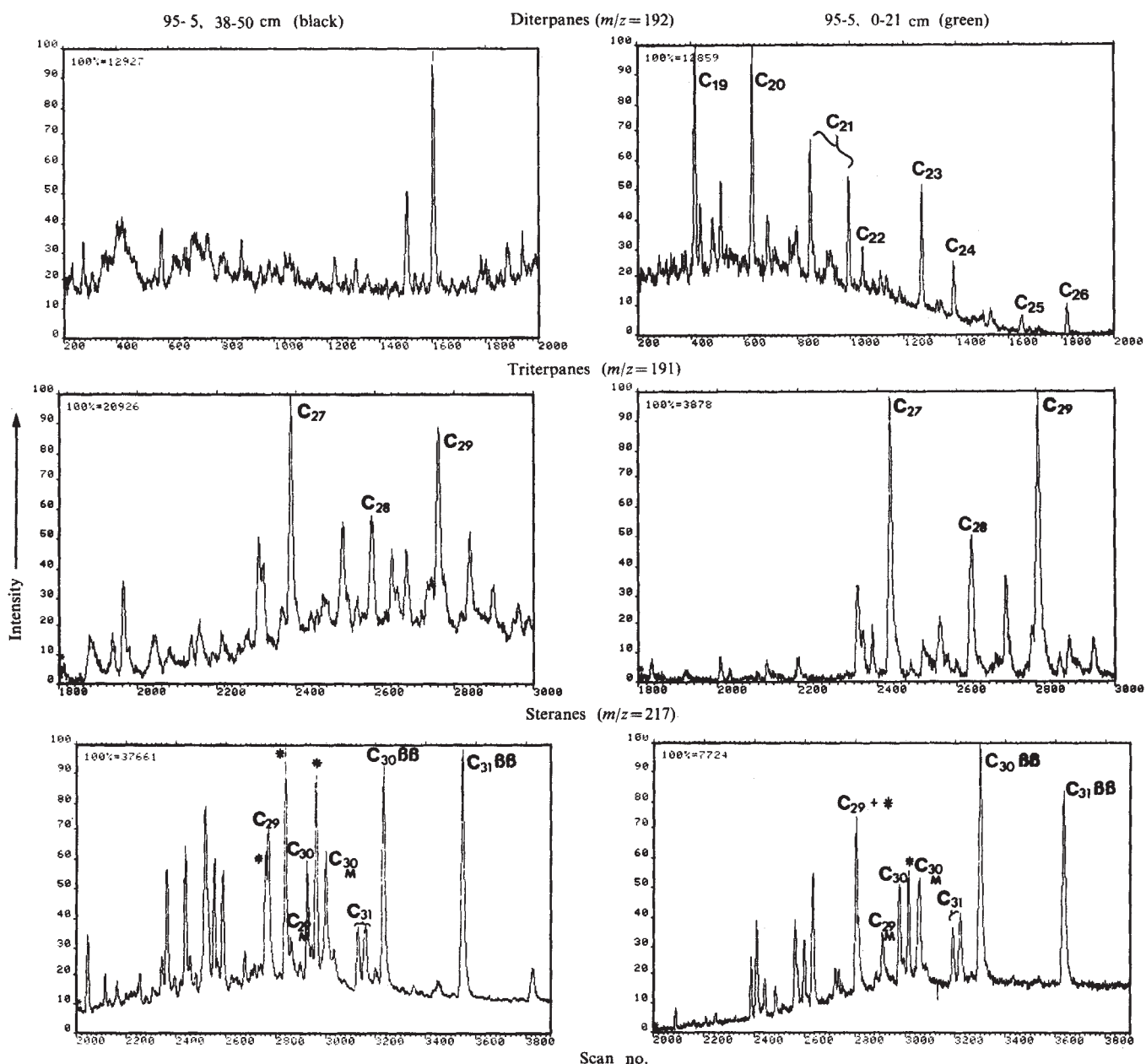


Fig. 3 GC-MS mass fragmentograms of steranes, triterpanes, and diterpanes extracted from black and green claystones from DSDP Hole 530A. The labelled steranes possess unaltered biological 5- α (H), 20- R configurations. The C₂₉ to C₃₁ triterpanes have 17- α (H), 21- β (H) configurations unless labelled as BB. Hopenes are indicated by *, and moretanes (17- β (H), 21- α (H) configurations) by M.

and alkanolic acids in all samples, suggest that a significant amount of these components were generated after deposition and, hence, may reflect diagenetic processes rather than selective preservation as suggested for the more labile compounds. Whereas both types of sample have relatively large amounts of long chain ($>C_{25}$) *n*-alkanes with a pronounced odd-carbon preference, suggesting terrigenous input, Fig. 2 shows that the black claystones also contain significant amounts of shorter chain *n*-alkanes and branched and cyclic hydrocarbons, which are not present in the green claystones^{14,15}. These aliphatic hydrocarbons and the larger quantities of C_{27} relative to C_{29} sterane, which are more evident in the black than in the green claystones (Fig. 3), suggest somewhat more aquatic input to the black claystone. In addition, the presence of diterpanes, predominantly C_{19} and C_{20} , in the green claystone, which are absent from the black claystone (Fig. 3) suggests the input of mostly terrigenous, microbially-altered material to the green claystone^{16,17}. The unidentified peaks in the m/z 191 trace of the black claystone probably represent tetracyclic terpanes of unknown origin. Hence, the hydrocarbon analysis agrees with the character of the organic matter intake suggested by the visual kerogen analysis.

The triterpane distributions are quite similar (Fig. 3): both types of sample exhibit abundant 17- β (H), 21- β (H)-hopanes which indicate a low degree of thermal maturity¹⁸. Both the triterpanes present and the occurrence of predominantly biological configurations (that is, 14- α (H), 17- α (H), 20-R)¹⁶ of the steranes support the low maturity indicated by vitrinite reflectance and spore colourations, and confirms that the recycled populations do not significantly contribute to the lipid content of these samples.

Comparison of distributions of *n*-alkanoic acids and *n*-alkanols reveals that the black shales contain relatively greater proportions of terrigenous than of marine components of these geolipid classes¹⁴. Because acids and alcohols are more subject to degradation than hydrocarbons, the difference in source character for these geolipid classes again suggests that selective preservation has occurred in the black shales.

The comparisons of organic matter contents confirm that black shales at DSDP Site 530 in the Angola Basin record periods of enhanced preservation of marine, as well as terrigenous, organic material which is superimposed on a low background of residual land-derived material¹⁵. Adjacent green claystones exhibit evidence of burrowing benthic animals and show that bottom waters must have been oxic in Cretaceous times^{3,19}. Because these Cenomanian sediments contain abundant turbidite sequences, the black shales probably represent episodes of downslope redeposition of sediments originally laid down in a midwater anoxic layer which impinged on the African continental slope^{19,20}. Such sediments, rich in organic matter from high marine productivity, must have been rapidly redeposited in the Angola Basin in order to preserve organic substances in an oxic environment. An important implication of this scenario is that deep-ocean black shales can result from local or regional events of short duration which do not involve global oceanic anoxia.

Received 27 February; accepted 13 September 1984.

1. Foresman, J. B. *Init. Rep. DSDP* 40, 557-567 (1978).
2. Bolli, H. M. *et al. Init. Rep. DSDP* 40, 357-455 (1978).
3. Hay, W. W. *et al. Bull. geol. Soc. Am.* 93, 1038-1050 (1982).
4. Summerhayes, C. P. & Masran, T. C. *Init. Rep. DSDP* 76, 469-480 (1983).
5. Dean, W. E. & Gardner, J. V. in *Nature and Origin and Cretaceous Carbon-rich Facies* (eds Schlanger, S.O. & Citas, M. B.) 55-78 (Academic, London, 1982).
6. Degens, E. T. & Mopper, K. in *Chemical Oceanography* 2nd edn (eds Riley, J. P. & Chester, R.) 5, 59-113 (Academic, London, 1976).
7. Demaison, G. J. & Moore, G. T. *Org. Geochem.* 2, 9-31 (1980).
8. Muller, P. J. *Geochim. cosmochim. Acta* 41, 756-776 (1977).
9. Nohara, M., Ishizuka, T. & Gieskes, J. M. *Init. DSDP* 75, 1951-1054 (1984).
10. Deroo, G., Herbin, J. P. & Huc, A. Y. *Init. Rep. DSDP* 75, 983-999 (1984).
11. Rullkötter, J., Mukhopadhyay, P. K. & Welte, D. H. *Init. Rep. DSDP* 75, 1069-1087 (1984).
12. Katz, B. L. *Init. Rep. DSDP* 75, 1031-1034 (1984).
13. Hatcher, P. G., Spiker, E. C., Szeverenyi, N. M. & Maciel, G. E. *Nature* 305, 498-501 (1983).
14. Meyers, P. A., Trull, T. W. & Kawka, O. E. *Init. Rep. DSDP* 75, 1009-1018 (1984).
15. Brassell, S. C. *Nature* 305, 1019-1030 (1983).
16. Simoneit, B. R. T. *Geochim. cosmochim. Acta* 41, 463-476 (1977).
17. Aquino Neto, F. R., Trendel, J. M., Restle, A., Connan, J. & Albrecht, P. A., in *Advances in Organic Geochemistry 1981* (ed. Björøy, M.) 659-667 (Wiley, Chichester, 1983).
18. Mackenzie, A. S. *et al. Geochim. cosmochim. Acta* 44, 1709-1721 (1980).
19. Dean, W. E., Arthur, M. A. & Stow, D. A. V. *Init. Rep. DSDP* 75, 819-844 (1984).
20. Thiede, J. & van Andel, T. H. *Earth planet. Sci. Lett.* 33, 301-309 (1977).

A model of estuarine circulation in the Pliocene Mediterranean based on new ostracod evidence

Dick van Harten

Geological Institute, University of Amsterdam, Nieuwe Prinsengracht 130, 1018 VZ Amsterdam, The Netherlands

The recent discovery of the psychrospheric ostracod *Agrenocythere pliocenica* (Seguenza) in Zanclean marls of eastern Crete¹ expands the reach of early Pliocene deep-oceanic circulation far into the eastern Mediterranean. Current models of Pliocene water-mass circulation, though allowing for the Atlantic psychrosphere to flow freely into the western basin, consider the eastern basin as being largely stagnant^{2,3}. I propose here a new model of estuarine circulation affecting both west and east Mediterranean. This circulation was associated with increased atmospheric humidity and driven by upwelling. The model is compatible with organic-rich sedimentation in parts of the eastern Mediterranean and does not imply an excessively deep or wide connection with the Atlantic.

The term 'psychrosphere'⁴ denotes, in an ecological sense, the cold (temperatures $<8-10^{\circ}\text{C}$), dense bottom-waters of the world ocean that are directly derived from the polar regions. The benthic ostracods inhabiting this environment include a group of highly characteristic genera, among them *Agrenocythere*, whose occurrence is virtually restricted to these waters. The depth at which these ostracods occur is generally $>500\text{ m}$ (ref. 2). In lower latitudes, they live at greater depth than in higher latitudes and upwelling regions⁵. *Agrenocythere* is typical of the upper psychrosphere: recent species are most likely to be found at depths of 1,000-1,500 m; the known depth range of the genus is 400-3,850 m (ref. 6).

In the modern Mediterranean, the temperature of deep bottom-water is $\sim 13^{\circ}\text{C}$ ⁷. Hence, recent faunas of this sea do not exhibit any typically psychrospheric ostracods. However, several such forms are recorded from Pliocene and earlier sediments in the Mediterranean region⁸. Figure 1 shows *A. pliocenica* found in Crete¹ which supports an earlier, doubtful identification of a single specimen in the Zanclean of DSDP Site 378 in the south Aegean⁹. All other spots where this species has been found lie much more to the west⁸⁻¹⁰; Sicily and southern peninsular Italy are the easternmost localities known so far. This new find eliminates any doubt that Pliocene psychrospheric circulation did press far into the east Mediterranean.

A. pliocenica first appeared in the Mediterranean in the *Globorotalia margaritae* zone (see Fig. 2 for the epochs, stages and biozones mentioned in this paper). In Italy, this species ranges up to the early Pleistocene⁹. In Crete, we have found it

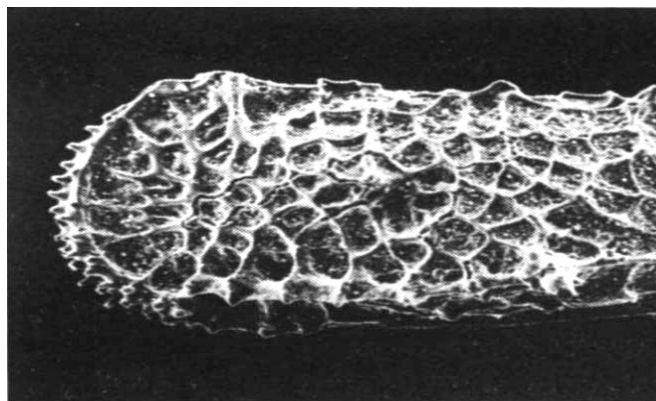


Fig. 1 *A. pliocenica* (Seguenza) from Zanclean marl (*G. margaritae* zone), Xerokambos, southeastern Crete, Greece. Left valve of a male. Length of specimen 1.4 mm.

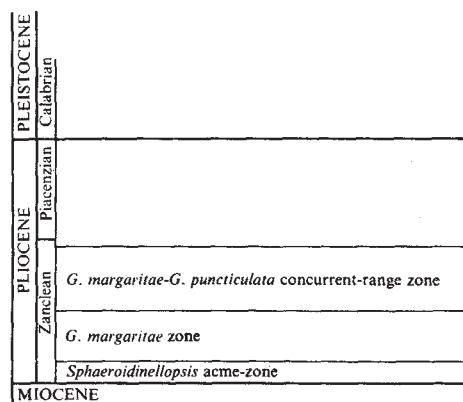


Fig. 2 Shows the relation of epochs, stages and biozones mentioned in the text. The Zanclean is intended to be synonymous with the Tabianian used by other authors to denote the earliest part of a bipartite Pliocene.

so far only in samples referable to the *G. margaritae* zone and possibly the *G. margaritae-G. puncticulata* concurrent-range zone. Psychrospheric conditions, although lasting into the Calabrian in Italy, do not seem to have reached the Zanclean-Piacenzian boundary in Crete.

At present, psychrospheric waters cannot enter the Mediterranean because the sill of the Strait of Gibraltar is too shallow (~320 m). To explain the psychrospheric influence in the Pliocene west Mediterranean, it has been suggested that a relatively wide and deep gateway lay between Spain and Morocco⁵. The sill of the connection with the ocean must indeed have been deeper than it is now, although it need not have been much deeper. Assuming dynamic circulation, a depth of 500–1,000 m must be sufficient to allow the uppermost, cold, deep water layers to pass. The width of the connection is difficult to judge, but it seems unnecessary to claim the existence of a true gateway.

The Sicilian Ridge, the shallow sill between Italy and Africa, divides the Mediterranean into a western and an eastern basin. The sill depth is now ~400 m. But the suggestion is that in the early Pliocene the Ridge was a high-standing barrier to deep circulation and thus encouraged stagnancy and the formation of organic-rich deposits in the eastern basin^{3,11}. This is inconsistent with the present evidence for psychrospheric circulation. The early Pliocene palaeodepth of the sill must at least have been equivalent to that of the strait into the ocean.

Participation of the psychrosphere in the circulation of the Pliocene Mediterranean can be inferred from the mere presence of benthic fauna. Otherwise, the deep water-masses would have stagnated and become uninhabitable through oxygen depletion. The circulation of the Mediterranean is now anti-estuarine (Fig. 3a). The high rate of evaporation, which exceeds the sum of river discharge and precipitation, leads to the formation of dense-surface waters that sink periodically to the bottom (downwelling). The densified Mediterranean water flows out into the Atlantic along the floor of the Strait of Gibraltar. Nearer the surface there is an Atlantic countercurrent which compensates for evaporation⁷. Assuming there is a sufficiently deep connection with the ocean, an estuarine pattern with deep water coming in and surface water going out is the simplest model to include the Atlantic psychrosphere in the Mediterranean circulation (Fig. 3b): upwelling can provide the drive for such a system in much the same way as downwelling operates the modern circulation.

The precondition for estuarine circulation is that the surface waters stay less dense than the deeper waters. This requires a stable pycnocline in the water-column in the form of either a halocline, thermocline, or thermohalocline. Evaporation tends to oppose this by making the surface water both saltier and cooler. An estuarine model therefore requires relatively high atmos-

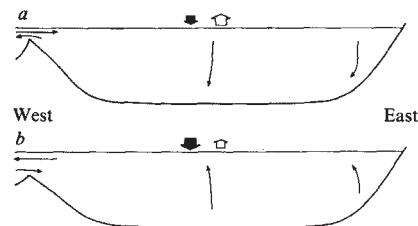


Fig. 3 Mediterranean watermass circulation: paired vertical arrows show schematically the relation of evaporation (open arrows) and fresh-water input by precipitation and river discharge (solid arrows). a, Modern-type anti-estuarine system: surface waters become locally dense and well down, eventually leaving the Mediterranean along the floor of the Strait of Gibraltar; surface currents bring in water from the Atlantic. b, Estuarine system as proposed for the Pliocene and earliest Pleistocene of southern Italy and part of the early Pliocene of Crete: Atlantic upper psychrospheric waters enter the Mediterranean under the influence of a pressure gradient; bottom-waters well up locally and mix with less dense water near the surface; surface currents are directed towards the Atlantic.

pheric humidity to suppress evaporation and promote precipitation. More precipitation leads to increased runoff which, in its turn, could contribute to a primary halocline in parts of the Mediterranean.

The presence of less dense waters in the Mediterranean water-column sets up a pressure gradient with the ocean. If the connection is deep enough, Atlantic psychrospheric waters will enter along this gradient. For these waters to participate in the Mediterranean circulation, upwelling is necessary: wind-driven upwelling most likely occurred in areas off the north coast under the influence of prevailing north winds. Such winds would have tended to move the upper water-layers to the west by Ekman transport. Thermohaline, density-driven upwelling, if at all significant, was probably limited to the easternmost parts where the ocean-derived deep waters had had maximum opportunity to absorb heat.

Estuarine Mediterranean circulation associated with a very warm and humid climate was suggested earlier for the *Sphaeroidinellopsis acme-zone*¹². Faunas of this zone show an unusual abundance of the planktonic foraminifer *Sphaeroidinellopsis*³. Similar non-spinose foraminifera with thick tests are known today to prefer deeper water¹³. The abundance of such forms in the earliest Pliocene supports the existence of a halocline with the upper layers too low in salt to be inhabited by marine plankton (S. R. Troelstra, personal communication).

From the occurrence of Taxodiaceae pollen it was recently concluded that humid climatic conditions and rainy summers prevailed in the early Pliocene of the northwest Mediterranean area^{14,15}. An important change, heralding an evolution towards modern seasonal fluctuations (dry summers and wet winters) occurred at the Zanclean-Piacenzian boundary (~3.2 Myr)¹⁵: summer aridity, inferred from the extinction of Taxodiaceae, stabilized diachronously over the Mediterranean region; stability was reached in the Piacenzian in southern France and in the Calabrian in Italy¹⁴. The 3.2 Myr climatic change is also reflected in the oxygen isotope curve of DSDP Site 132 in the Tyrrhenian where it was tentatively associated with increased evaporation relative to precipitation¹⁶.

Sedimentary manifestations of the proposed upwelling might be found in several organic-rich layers or sapropels that have been reported from the early Pliocene of the east Mediterranean, especially at DSDP Site 378 (ref. 17). These sapropels tend to differ from those of the Messinian, late Pliocene, and Quaternary by being relatively low in organic material and burrowed and by containing benthic fauna. They might reflect boosted primary productivity associated with the upwelling of nutrient-rich water from the bottom, rather than complete anoxia due to basin-wide stagnancy¹¹.

Psychrospheric bottom conditions apparently lasted much

longer in Italy than in Crete. At present we can only speculate as to their cause. Considering the similar diachronous summer-aridity stabilization in Italy and France, differential climatic effects may seem a probable cause but direct palynological evidence for this from the east Mediterranean is still lacking. Alternatively, changes in the sea-bottom topography, for instance a rising of parts of the Sicilian Ridge, may have obstructed some deep circulation routes while leaving others unimpaired.

The model outlined here needs further critical testing. Topics to be investigated especially include the climatic history of the east Mediterranean and the relation of Mediterranean organic-rich sediments to upwelling.

I thank Dr Richard H. Benson and Mr Jeroen M. Peters for critical reading of the manuscript and comment. The scanning electron micrograph was made in cooperation with 'Werkgroep SEM', University of Amsterdam.

Received 8 June; accepted 13 September 1984.

- Peters, J. M., Troelstra, S. R. & van Harten, D. J. *geol. Soc. Lond.* (in the press).
- Benson, R. H. in *The Mediterranean Sea: a Natural Sedimentation Laboratory* (ed. Stanley, D. J.) 63-73 (Dowden, Stroudsburg, 1972).
- Cita, M. B. & Ryan, W. B. F. *Init. Rep. DSDP* 13, 1405-1415 (1973).
- Bruun, A. F. *Mem. geol. Soc. Am.* 67 (1), 641-672 (1957).
- Benson, R. H. & Sylvester-Bradley, P. C. *Bull. Cent. Rech. Pau* 5 suppl. 63-91 (1971).
- Benson, R. H. *Smithson. Contr. Paleobiol.* 12, 1-138 (1972).
- Lacombe, H. & Tchernia, P. in *The Mediterranean Sea: a Natural Sedimentation Laboratory* (ed. Stanley, D. J.) 25-36 (Dowden, Stroudsburg, 1972).
- Benson, R. H. *Init. Rep. DSDP* 42 (i), 777-787 (1978).
- Colalongo, M. L. & Pasini, G. *Boll. Soc. Paleont. Ital.* 19, 44-126 (1980).
- Carbonnel, G. & Ballesio, R. *Docum. Lab. géol. Lyon* 85, 1-113 (1982).
- Cita, M. B. & Grignani, D. in *Nature and Origin of Cretaceous Carbon-Rich Facies* (eds Schlanger, S. O. & Cita, M. B.) 165-196 (Academic, New York, 1982).
- van Gorsel, J. T. & Troelstra, S. R. *Géol. Méditerran.* 7, 127-134 (1980).
- Bé, A. W. H. & Tolderlund, D. S. in *Micropalaeontology of Oceans* (eds Funnell, B. M. & Riedel, W. R.) 105-149 (Cambridge University Press, London, 1971).
- Suc, J.-P. & Zagwijn, W. D. *Boreas* 12, 153-166 (1983).
- Suc, J.-P. *Nature* 307, 429-432 (1984).
- Keigwin, Jr., L. D. & Thunell, R. C. *Nature* 282, 294-296 (1979).
- Kidd, R. B., Cita, M. B. & Ryan, W. B. F. *Init. Rep. DSDP* 42 (i), 421-443 (1978).

Potassium-selective single channels in guard cell protoplasts of *Vicia faba*

J. I. Schroeder*, R. Hedrich† & J. M. Fernandez*

* Department of Membrane Biophysics, Max-Planck-Institut für biophysikalische Chemie, D-3400 Göttingen, FRG

† Pflanzenphysiologisches Institut, Universität Göttingen, D-3400 Göttingen, FRG

In plants, stomata control the gaseous exchange between the intercellular spaces of the leaf and the atmosphere. Fluxes of ions, in particular potassium fluxes, across the membranes of the guard cells produce changes in turgor of the guard cells which in turn result in the opening and closing of the stomatal pore^{1,2}. The molecular mechanisms involved in the uptake or release of ions in guard cells are poorly understood². Cell-free membrane patches of the plasmalemma, isolated following patch-clamp techniques³⁻⁵, allow, for the first time in higher plant cells, the separation of the electrical properties of the plasmalemma from those of the tonoplast. We have applied these techniques to study the properties of single-ion channels in the plasmalemma of guard cell protoplasts of *Vicia faba* (broad bean). Predominantly a cation-selective channel was observed, which showed a high selectivity for K⁺, with a permeability ratio P_{K^+}/P_{Na^+} of 11:1 and a single-channel conductance of 37 pS ($=37 \times 10^{-12} \Omega^{-1}$) in symmetrical 225 mM KCl solutions. We estimate that this K⁺ channel contributes significantly to the uptake and release of K⁺ by guard cells during stomatal movement.

Guard cell protoplasts were isolated from the lower epidermis of 3-week-old bean leaves (*Vicia faba* Weisskeimige Hang-down). Epidermal strips were exposed for 60-90 min to 2% cellulase (Rohament CT; Röhm), 0.5% pectinase (Rohament P; Röhm), 0.4 M mannitol and 1 mM CaCl₂. Purified guard cell protoplasts were obtained after 2 h, including the time required

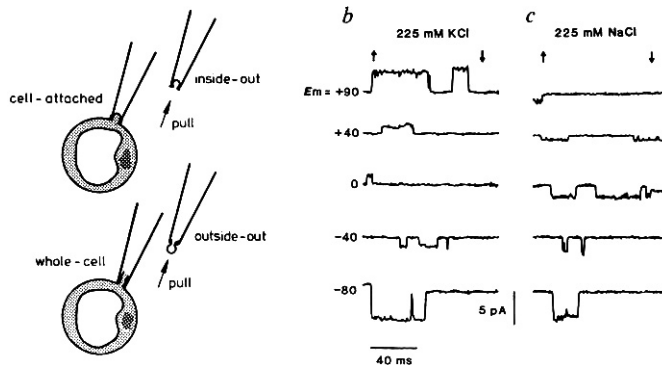


Fig. 1 *a*, Schematic diagram of all four patch-clamp recording configurations³, applied to guard cell protoplasts. Upper panel: high-resistance seals between the pipette and the membrane are established when the fire-polished electrode tip is pressed against the surface of the membrane and suction is applied to the inside of the pipette (cell-attached recording mode). The mechanical stability of the seals allows small membrane fragments to be torn off the cell by drawing the pipette away from the cell. This procedure leaves an intact patch of membrane spanning the pipette tip, resulting in an inside-out patch. The cytoplasmic face of the plasmalemma is exposed to the bath solution. Lower panel: an outside-out patch is obtained when the membrane patch is disrupted by a pulse of suction before pulling off the pipette. The correct sidedness of these patch configurations has been confirmed in numerous experiments on animal cells with channels of known asymmetry³. *b*, *c*, Recordings of K⁺-selective channel currents in an inside-out membrane patch. Discrete open-closed state transitions of single-channel currents can be seen. The membrane was held at $E_m = +40$ mV and stepped by voltage pulses to potentials ranging from +90 mV to -80 mV. The voltage pulse duration was 100 ms. In all records the pulse starts at the upward-pointing arrow and stops at the downward-pointing arrow. Membrane potentials are specified as the potential on the cytoplasmic side of the membrane relative to that on the extracellular side of the membrane during the voltage pulse. Leakage and capacitive currents were subtracted at each potential, using records in which no channel activity was visible. *b*, In symmetrical high K⁺ solutions (225 mM KCl) the single-channel currents are outward at +40 mV and +90 mV (upward deflections of trace). Currents are inward at -40 mV and -80 mV. Note the increase in noise when the channels are open¹⁹. The reversal potential of channel currents is close to 0 mV. Note that on the 0 mV and -80 mV traces, a channel opened before the voltage pulse was turned on and the membrane potential was still +40 mV. When the pulse was turned on, the single-channel current vanished in the 0 mV trace. *c*, After K⁺ was replaced by Na⁺ in the bathing medium (facing the cytoplasmic face of the membrane), the reversal potential shifted to almost +80 mV. A channel was open when the pulse was turned on in the +90 mV trace. All currents are inward from +40 to -80 mV. These data indicate a high selectivity of the channel for K⁺ over Na⁺. Solutions: extracellular (pipette; in mM): 225 KCl, 2 MgCl₂, 1 CaCl₂, 3 KOH, 10 HEPES, 80 D-mannitol, pH 7.0; intracellular (bath): in *b*, same as pipette; in *c*, 225 NaCl, 2 MgCl₂, 1 CaCl₂, 3 NaOH, 10 HEPES, 80 D-mannitol, pH 7.0. All solutions were adjusted to a final osmolarity of $\pi = 500$ mmol kg⁻¹, by adding D-mannitol. Osmolarities were measured with a vapour pressure osmometer (Wescor 5100 C). Temperature 22-25 °C. Control of membrane potential and the measurement of currents were performed with an EPC-7 patch-clamp (List Electronic) and low-pass-filtered at 1 kHz with an 8-pole Bessel characteristic. Data were sampled and stored on a PDP-11/23 minicomputer operating on-line.

for incubation, filtration through a 20 μ m nylon net and two centrifugation steps (60g and 45g). They were suspended in 0.5 M mannitol, 1 mM CaCl₂, stored on ice and used for experiments within 6 h after isolation.

High-resistance seals (>10 G Ω) between the fire-polished tip of the recording pipette and the plasmalemma were established, allowing high-resolution recordings of single-channel currents³. The best results for seal formation were obtained within the first 2 h after protoplast isolation and using small pipettes (tip resistance >3 M Ω)⁶. Techniques well established in animal

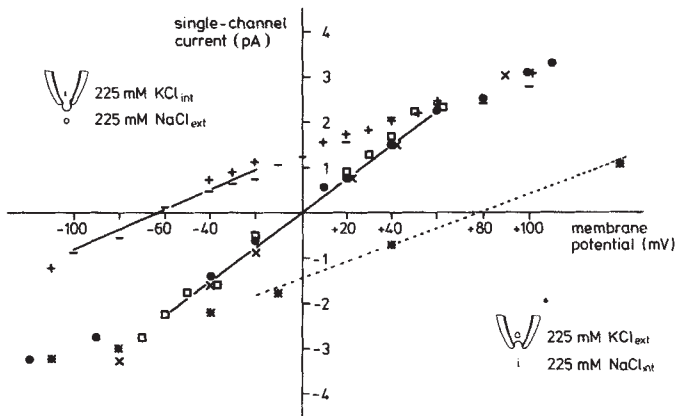


Fig. 2 Amplitude of single-channel currents as a function of membrane potential in three cell-free membrane patches (two outside-out patches, one inside-out patch). The three curves were determined by least-squares regression analysis of data in the given intervals. Outward currents have a positive sign. Middle curve: ●, □ × (× is inside-out patch), data from patches in symmetrical 225 mM KCl solutions. The I - V relationship is approximately ohmic between -60 mV and $+60$ mV. The slope reveals a conductance of 37 ± 4 pS. Left curve and top left insert: The insert shows an outside-out patch spanning the pipette tip. +, -, Data from the same two outside-out patches after the bath solution was changed to 225 mM NaCl as in Fig. 1c. The reversal potential of single-channel currents shifted to -61 mV (see text). The slope conductance of single-channel currents was reduced to 22 ± 2 pS (-100 mV to -20 mV). Right curve and bottom right insert: *, Data from inside-out membrane patch after the bath solution (facing the cytoplasmic membrane side) was changed to 225 mM NaCl. The reversal potential shifted to $\sim +78$ mV, corresponding to a permeability ratio P_{K^+}/P_{Na^+} of 21:1. The single-channel conductance was fitted to 19 pS between -10 mV and $+140$ mV. Temperature 22 – 25°C .

physiology were used to produce cell-free membrane patches with either the cytoplasmic membrane face (inside-out patch) or the external membrane face (outside-out patch) exposed to the bathing medium^{3,4} (Fig. 1a). Single-channel current could be studied in these configurations, while defining ionic conditions on both sides of the membrane (Fig. 1b, c).

The most frequent ion channel observed was a K^+ -selective channel (Fig. 1b, c) which appeared in 90% of the patches (patch area 2 – $10 \mu\text{m}^2$)⁶. The conductance of this channel depended on the K^+ concentration: in solutions containing 30 mM K^+ on the extracellular and 228 mM K^+ on the cytoplasmic side of the membrane, the single-channel conductance ranged from 20 to 27 pS. The direction of single-channel currents reversed at membrane potentials of $E_m = -50$ mV (reversal potential), corresponding to the Nernst potential for K^+ (not illustrated). In symmetrical high K^+ solutions (both pipette and bath containing 225 mM KCl), the single-channel conductance was 37 ± 4 pS ($\pm$ s.d., $n = 3$) (Fig. 2, middle curve).

The permeability ratio of K^+ to Na^+ was determined by measuring the bi-ionic potential^{7,8}. K^+ in the bathing medium was replaced with Na^+ by bath perfusion (Fig. 1b, 225 mM NaCl); this shifted the reversal potential of single-channel currents from 0 mV to approximately -61 mV in outside-out patches. The single-channel current-voltage (I - V) relationship was approximately linear between -100 and -20 mV with a conductance reduced to 22 pS (Fig. 2, left curve). These data indicate a permeability ratio for K^+ to Na^+ (P_{K^+}/P_{Na^+}) of 11:1, after correcting for ionic activities⁹. When K^+ was replaced by Na^+ on the cytoplasmic side of the membrane in an inside-out patch, the reversal potential of single-channel currents shifted to $\sim +78$ mV, indicating an even higher selectivity for K^+ over Na^+ (Fig. 2, right curve).

The kinetics of opening and closing of the K^+ -selective channel were analysed in a patch in which apparently only one

channel was present¹⁰. The mean lifetime of an open channel at $E_m = -60$ mV was 7 ms ($n = 76$) and 10 ms at $E_m = +60$ mV ($n = 60$). The mean closed time of the channel was ~ 13 ms at both -60 mV and $+60$ mV ($n = 103$ and 72 , respectively). These results indicate that the gating mechanism, responsible for the opening and closing of the K^+ -selective channel, is not significantly voltage-dependent and that the channel is open approximately one-third of the time.

The appearance of channels is not an artefact of patch isolation, as the same K^+ -selective channel was observed in the plasmalemma of intact guard cell protoplasts (cell-attached recording mode)³ (Fig. 1a), as judged by the slopes and reversal potentials of the I - V relationships.

K^+ -salts are the predominant osmotically active compounds responsible for changes in the stomatal aperture in over 50 species^{1,2}. Several authors have described a selectivity for K^+ over Na^+ in conjunction with stomatal movement^{11–13}. Guard cells of *V. faba* accumulate between 100 and 300 mM K^+ when the stomata change from the closed to the open state². Furthermore, the relationship between the resting potential of guard cells and external K^+ concentration has been shown to be approximately nernstian^{14,15}. The mean estimated flux of K^+ into guard cells of *V. faba*, determined by several authors², is $10 \text{ pmol s}^{-1} \text{ per cm}^2$ of plasmalemma surface. These K^+ fluxes are probably associated with membrane potential shifts of up to 60 mV, caused by effectors on stomatal aperture, such as light, CO_2 and abscisic acid^{16–18}. If the membrane potential is displaced 30 mV from the K^+ equilibrium potential, we predict an initial flux through the K^+ -selective channels in a single guard cell of $20 \text{ pmol s}^{-1} \text{ cm}^{-2}$, corresponding to a change in the cell's K^+ concentration of $60 \mu\text{M s}^{-1}$. In this calculation we assumed a channel density of one channel per $15 \mu\text{m}^2$, a guard cell volume of 5 pl , a single-channel conductance of 30 pS, and that the channel is open one-third of the time. This estimate demonstrates that a significant part of K^+ uptake and release by guard cells can be accounted for by passive transport through the K^+ -selective channel studied here.

In addition to the predominant K^+ channel described, several other types of ion channels were seen but have not yet been characterized properly; patch-clamp studies, as used here, may elucidate their physiological function.

We thank Dr E. Neher, Professor K. Raschke, Dr F. J. Sigworth and Professor D. Gradmann for discussions and comments on the manuscript; and U. Brill, W. Lahr, U. Lohmann, M. Odenthal, W. Pagel, A. Sauer, H. Terasa and F. Tiefenau for peeling bean leaves for protoplast preparation. J.M.F. is an Alexander von Humboldt fellow and J.I.S. was supported by the Deutsche Forschungsgemeinschaft.

Received 23 July; accepted 28 September 1984.

1. Raschke, K. in *Encyclopedia of Plant Physiology* Vol. 7 (eds Haupt, W. & Feinlieb, M. E.) 384–441 (Springer, Berlin, 1979).
2. Outlaw, W. H. *Physiologia Pl.* **59**, 302–311 (1983).
3. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1980).
4. Horn, R. & Patlak, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6930–6934 (1980).
5. Sakmann, B. & Neher, E. (eds) *Single Channel Recording* (Plenum, New York, 1983).
6. Sakmann, B. & Neher, E. (eds) in *Single Channel Recording*, 37–51 (Plenum, New York, 1983).
7. Läuger, P. *Biochim. biophys. Acta* **311**, 432–436 (1973).
8. Coronado, R., Rosenberg, R. L. & Miller, C. J. *gen. Physiol.* **76**, 425–446 (1980).
9. Robinson, R. A. & Stokes, R. H. *Electrolyte Solutions* (Academic, New York, 1955).
10. Colquhoun, D. & Hawkes, A. G. in *Single Channel Recording* (eds Sakmann, B. & Neher, E.) 135–174 (Plenum, New York, 1983).
11. Imamura, S. *Jap. J. Bot.* **12**, 251–346 (1943).
12. Fischer, R. A. & Hsiao, T. C. *Pl. Physiol.* **43**, 1958–1968 (1968).
13. MacRobbie, E. A. C. *J. exp. Bot.* **34**, 1695–1710 (1983).
14. Cork, R. J. in *Membrane Permeability: Experiments and Models* (ed. Bretag, A. H.) 57–60 (Techsearch Inc., Adelaide, 1983).
15. Edwards, E. & Bowling, J. F., *J. exp. Bot.* **35** (153), 562–567 (1984).
16. Zeiger, E., Moody, W., Hepler, P. & Varela, F. *Nature* **270**, 270–271 (1977).
17. Ishikawa, H., Aizawa, H., Kishira, H., Ogawa, T. & Sakata, M. *Pl. Cell Physiol., Tokyo* **24**, 769–772 (1983).
18. Gotow, K., Kondo, N. & Syono, K. *Pl. Cell Physiol., Tokyo* **23** (6), 1063–1070 (1982).
19. Sigworth, F. J. in *Single Channel Recording* (eds Sakmann, B. & Neher, E.) 306–314 (Plenum, New York, 1983).

Benzodiazepines antagonize cholecystokinin-induced activation of rat hippocampal neurones

Jacques Bradwejn & Claude de Montigny*

Neurosciences Research Center, University of Montreal, Montreal, Quebec, Canada H3C 3J7

* To whom correspondence should be addressed

Cholecystokinin (CCK) is a neuropeptide present in the mammalian central nervous system (CNS)¹. In all species studied so far, the highest concentrations of this neuropeptide have been found in the cerebral cortex, the amygdala and the hippocampus²⁻⁵. Five molecular forms of CCK having 39, 33, 13, 8 and 4 amino acid residues have been identified in the CNS, the sulphated octapeptide (CCK₈) being the most abundant form detected^{2,6-8}. Specific CCK binding sites have been demonstrated in the rat, guinea pig and human brain^{9,10}. CCK₈, applied by microiontophoresis to deep cortical neurones and hippocampal pyramidal neurones, has a powerful excitatory effect, whereas the non-sulphated CCK octapeptide has no such effect on these neurones¹¹⁻¹⁶. Low doses of benzodiazepines depress the spontaneous activity of hippocampal pyramidal neurones¹⁷⁻²⁰. We report here that benzodiazepines at very low doses antagonize selectively the CCK₈-induced activation of rat hippocampal pyramidal neurones. This antagonistic action might be involved in the anxiolytic effect of these drugs.

Male Sprague-Dawley rats (200-300 g) were anaesthetized with urethane (1.25 g per kg intraperitoneal) and mounted in a stereotaxic apparatus. Five-barrelled micropipettes were prepared in the conventional manner²¹. The central barrel, filled with 2 M NaCl saturated with Fast Green FCF, was used for extracellular recording of unitary activity. Recordings were obtained from pyramidal neurones of the CA1 and CA3 regions of the dorsal hippocampus. The signal, amplified and displayed in the usual manner, was fed into a differential amplitude

discriminator, generating square pulses from which integrated firing rate histograms were obtained. The side barrels were used for microiontophoresis of the following substances: CCK₈ (10 μ M in 0.2 M NaCl, pH 5); acetylcholine chloride (20 mM, pH 4); flurazepam-2HCl (20 mM, pH 4); chlordiazepoxide-HCl (10 mM in 20 mM NaCl, pH 3); Met-enkephalin (0.5 mM in 0.2 M NaCl and 0.01% bovine serum albumin, pH 4.6); L-glutamate (50 mM in 50 mM NaCl, pH 8.0); L-aspartate (50 mM in 50 mM NaCl, pH 8.0). Three side barrels were used for ejection of the test substances and the fourth one, filled with 2 M NaCl, was used for current balancing. Recording sites were marked by passing a -27 μ A current through the central barrel for 20 min, to make a Fast Green deposit, the location of which was verified histologically.

Intravenous (i.v.) administration of flurazepam, lorazepam or diazepam during the activation of 18 hippocampal pyramidal neurones by microiontophoretically applied CCK₈ reversed this activation, whereas the excitatory effects of applications of acetylcholine (ACh) (14 neurones) or Met-enkephalin (7 neurones) were not affected (Fig. 1a). The effect of the three benzodiazepines was the same in the CA1 and CA3 regions. From the dose-response curves of the effect of i.v. lorazepam and diazepam on CCK₈-induced activation (Fig. 2), we calculated the 50% effective dose (ED₅₀) to be 32.6 ± 2.8 (mean $\pm$ s.e.m.) and 106.5 ± 18.6 μ g per kg for lorazepam and diazepam, respectively.

To verify further the selectivity of benzodiazepines for CCK₈ and to determine whether they were reversing CCK₈-induced activation by a direct action on the pyramidal neurone recorded, flurazepam or chlordiazepoxide, two water-soluble benzodiazepines, were applied microiontophoretically during the activation of CA1 or CA3 hippocampal neurones by applications of CCK₈ (40 neurones), Met-enkephalin (14 neurones), glutamate (8 neurones) and aspartate (8 neurones). Microiontophoretic applications of either benzodiazepine reversed the CCK₈-induced activation of pyramidal neurones but not that produced by ACh, Met-enkephalin, glutamate or aspartate (Fig. 1b-d).

As a preliminary investigation of the specificity of this antagonism for benzodiazepines, haloperidol (5 mg per kg), an

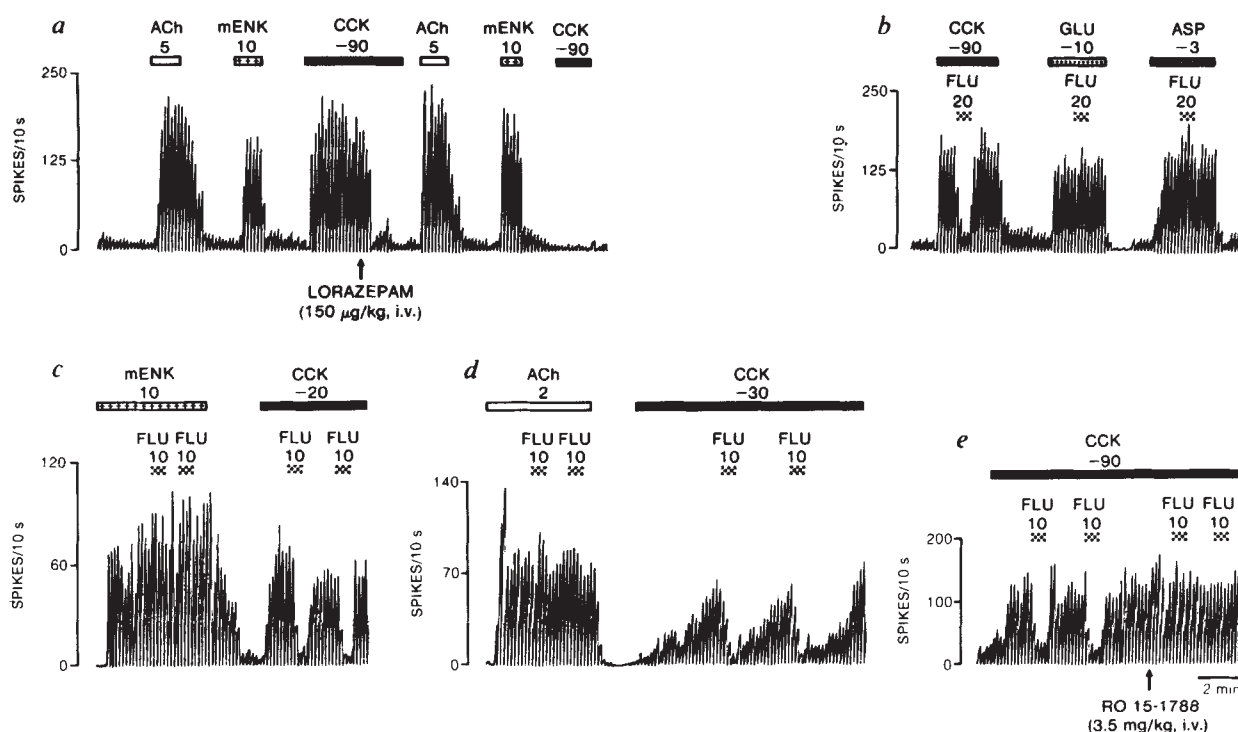


Fig. 1 Histograms of integrated firing rate for pyramidal neurones of the dorsal hippocampus CA3 region. Bars indicate duration of microiontophoretic applications, for which the current is given in nA. The time scale in *e* applies to all the traces. CCK, cholecystokinin-8; ACh, acetylcholine; mENK, Met-enkephalin; FLU, flurazepam; GLU, glutamate; ASP, aspartate.

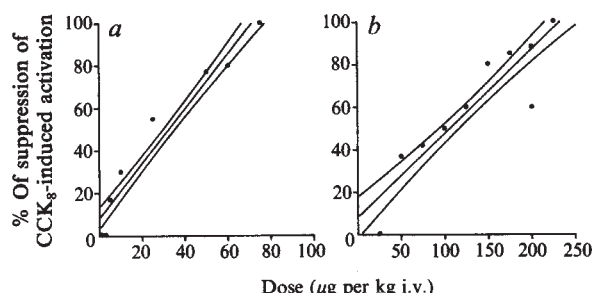


Fig. 2 Dose-response curve ($\pm$ s.e.m.) of the degree of suppression by i.v. lorazepam (a) and diazepam (b) of the activation of hippocampal pyramidal neurones induced by microiontophoretic application of CCK₈. $r = 0.94$, $P < 0.001$ for lorazepam; $r = 0.82$, $P < 0.001$ for diazepam.

antipsychotic drug, phenobarbital (10 mg per kg), an anticonvulsant barbiturate, and meprobamate (20 mg per kg), a non-benzodiazepine anxiolytic drug, were administered i.v. during the activation of CA3 pyramidal neurones by CCK₈. None of the three non-benzodiazepine drugs affected the CCK₈-induced activation in the 10 experiments performed.

To ascertain that benzodiazepines antagonize CCK₈-induced activation via their specific receptors^{22,23}, we used the imidazodiazepine Ro 15-1788, a benzodiazepine antagonist²⁴. Ro 15-1788 (3.5 mg per kg i.v.) completely blocked the effects of microiontophoretic applications of flurazepam and chlordiazepoxide and of the i.v. administration of lorazepam (100 µg per kg) on CCK₈-induced activation in all six rats tested (Fig. 1d). Moreover, in four other experiments, Ro 15-1788 (3.5 mg per kg i.v.) completely reversed the antagonism of the CCK₈-induced activation produced by microiontophoretic application of flurazepam and i.v. lorazepam (not shown).

This is the first electrophysiological demonstration of the antagonism of the action of a neuropeptide on CNS neurones by benzodiazepines. The present data do not allow any conclusion regarding the molecular mechanism underlying the antagonism by benzodiazepines of the neuronal activation by CCK₈. The selectivity of the action of benzodiazepines for CCK₈-induced activation is indicated by their inability to affect, when administered either intravenously or microiontophoretically, the excitatory effects of another peptide, Met-enkephalin, of ACh and of two putative neurotransmitter amino acids, glutamate and aspartate. Benzodiazepines have been shown to antagonize the excitatory action of kainate on CA1 hippocampal pyramidal neurones²⁵. However, doses more than 30 times greater than those used here are required to reduce kainate-induced activation²⁵; moreover, the activation of CA3 pyramidal neurones by kainate is not antagonized by benzodiazepines²⁵ whereas the excitatory action of CCK₈ is blocked equally well by benzodiazepines in the CA1 and CA3 regions.

The activity of all four benzodiazepines tested in the present study (flurazepam, diazepam, lorazepam and chlordiazepoxide) suggests that the ability to antagonize CCK₈-induced activation is a property common to all drugs of this class. The inactivity of the three anxiolytic non-benzodiazepine drugs tested (haloperidol, phenobarbital and meprobamate) constitutes a preliminary indication that this property might be specific to benzodiazepines. That the antagonism of CCK₈-induced activation might be a property common to, and specific for, benzodiazepines is fully consistent with the prevention and the reversal by Ro 15-1788, a specific 'neuronal' benzodiazepine antagonist, of the effect of benzodiazepines on CCK₈-induced activation²⁶. As there is strong evidence that the activation of the neuronal type of benzodiazepine receptors mediates the anxiolytic effect of benzodiazepines^{27,28}, this suggests that the antagonism of CCK₈ action by these drugs might be involved in their therapeutic effect. The extreme potency of benzodiazepines in blocking CCK₈-induced activation is consistent with this hypothesis.

It is indeed striking that the ED₅₀ for lorazepam (32 µg per kg) and diazepam (106 µg per kg) found here are very close to the most widely prescribed single doses of 1 mg lorazepam and 5 mg diazepam (equivalent to 15–20 µg per kg and 75–100 µg per kg, respectively, for an adult of average weight) in the treatment of anxiety. Moreover, the ratio of the potencies of lorazepam and diazepam found here is very close to that of their clinical anxiolytic activities. Finally, the high density of CCK₈ in the cerebral cortex and in limbic regions such as the hippocampus and the amygdala^{2–4} makes this neuropeptide a suitable candidate for mediating the selective anxiolytic effects of benzodiazepines. As benzodiazepine receptors are coupled with γ -aminobutyric acid (GABA) receptors²⁹, there is a need to investigate the involvement of GABA in mediating the selective suppression by benzodiazepines of CCK₈-induced activation.

We thank W. Haefely for providing us with Ro 15-1788, G. Dubé for typing, and D. Cyr and G. Filosi for preparing the illustrations. C. de M. is a Scientist of the MRC of Canada and J.B. a Fellow of the Fonds de la Recherche en Santé du Québec.

Received 11 June; accepted 10 September 1984.

- Morley, J. E. *Life Sci.* **30**, 479–494 (1982).
- Larsson, L. I. & Rehfeld, J. F. *Brain Res.* **165**, 201–218 (1979).
- Rehfeld, J. F. *J. biol. Chem.* **253**, 4022–4030 (1978).
- Geola, F. L. et al. *J. clin. Endocr. Metab.* **53**, 270–275 (1981).
- Vanderheagen, J.-L. *INSERM Eur. Symp.* **5** (eds Dumont, J. E. & Nurez, J.) 149–152 (Elsevier, Amsterdam, 1981).
- Goltermann, N. R., Rehfeld, J. F. & Roigaard-Petersen, H. *J. biol. Chem.* **255**, 6181–6185 (1980).
- Goltermann, N. R., Rehfeld, J. F. & Roigaard-Petersen, H. *J. Neurochem.* **35**, 479–483 (1980).
- Goltermann, N. R., Stengaard-Pedersen, K., Rehfeld, J. F. & Christensen, N. J. *J. Neurochem.* **36**, 959–965 (1981).
- Hays, S. E., Beinfeld, M. C., Jensen, R. L., Goodwin, F. D. & Paul, S. M. *Neuropeptides* **1**, 53–62 (1980).
- Innis, R. B. & Snyder, S. H. *Eur. J. Pharmac.* **65**, 123–124 (1980).
- Oomura, Y., Orita, M., Kita, H., Ishibashi, S. & Okajima, T. in *Iontophoresis and Transmitter Mechanisms in the Mammalian Central Nervous System* (eds Ryall, R. W. & Kelly, J. S.) 120–123 (Elsevier, Amsterdam, 1978).
- Phillis, J. W. & Kirkpatrick, J. R. *Can. J. Physiol. Pharmac.* **58**, 337–340 (1980).
- Dodd, J. & Kelly, J. S. *Brain Res.* **205**, 337–350 (1981).
- Kelly, J. S. & Dood, J. in *Neurosecretion and Brain Peptides* (eds Martin, J. S., Reichlin, S. & Bick, K. L.) 133–144 (Raven, New York, 1981).
- Bunney, B. S., Grace, A. A., Hommer, D. W. & Skirboll, L. R. in *Regulatory Peptides: From Molecular Biology to Function* (eds Costa, R. & Trabucchi, M.) 424–436 (Raven, New York, 1982).
- Skirboll, L. R. et al. *Neuroscience* **6**, 2111–2124 (1981).
- Adamec, R. E., McNaughton, B., Racine, R. & Livingston, K. E. *Epilepsia* **22**, 205–215 (1981).
- Lee, H. K., Dunwiddie, T. W. & Hoffer, B. J. *Naunyn-Schmiedeberg Archs Pharmac.* **309**, 131–136 (1979).
- Tsuchiya, T. & Fukushima, H. *Eur. J. Pharmac.* **48**, 421–424 (1978).
- Wolf, P. & Haas, H. L. *Naunyn-Schmiedeberg Archs Pharmac.* **299**, 211–218 (1977).
- Haigler, H. J. & Aghajanian, G. K. *J. Pharmac. exp. Ther.* **188**, 688 (1974).
- Squires, R. F. & Braestrup, C. *Nature* **266**, 732–734 (1977).
- Mohler, H. & Okada, T. *Science* **198**, 849–851 (1977).
- Hunkeler, W. et al. *Nature* **209**, 514–516 (1981).
- de Bonnel, G. & de Montigny, C. *Eur. J. Pharmac.* **93**, 45–54 (1983).
- Mohler, H., Burkard, W. P., Keller, H. H., Richards, V. G. & Haefely, W. J. *Neurochem.* **33**, 714–722 (1981).
- Braestrup, C. & Squires, R. F. *Eur. J. Pharmac.* **48**, 263–270 (1978).
- Lippa, A. S. et al. *Pharmac. Biochem. Behav.* **9**, 853–856 (1978).
- Haefely, W. & Polc, P. in *Anxiolytics: Neurochemical, Behavioral and Clinical Perspectives* (eds Malik, J. B., Enna, S. J. & Yamamura, H. I.) 113–145 (Raven, New York, 1983).

Immunohistochemical localization of endogenous nerve growth factor

R. A. Rush

Department of Human Physiology and Centre for Neuroscience, Flinders University of South Australia, Bedford Park, South Australia 5042, Australia

Nerve growth factor (NGF) has been proposed as a trophic molecule essential for the development of sympathetic and primary sensory neurones^{1,2}. In newborn mice and rats, administration of nerve growth factor results in an increase in the number of surviving neurones^{3,4}, whereas administration of antiserum to NGF decreases neuronal survival^{5,6}. Thus it has been proposed that the factor is produced and secreted by the relevant target tissues to provide trophic support for the ingrowing nerves^{1,7}. The site of synthesis of nerve growth factor is still unknown, and it has been emphasized

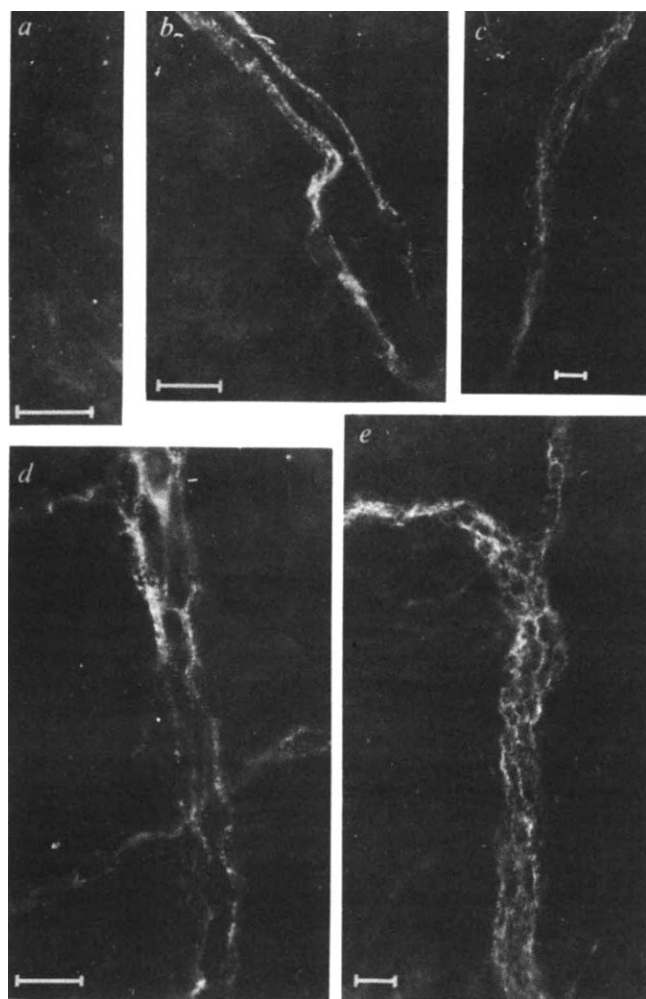


Fig. 1 Anti-NGF immunohistochemistry in the rat iris. *a*, The absence of stain in the non-cultured iris; *b-e*, the presence of stain in cells clustered into bundles following 4 or 6 days culture. Irides were cultured for 4 (*b, d*) or 6 (*c, e*) days in Eagle's basal medium containing 10% fetal calf serum and exposed to a humidified 5% CO₂/95% air atmosphere. After incubation the irides were removed, freeze thawed and stretched on polyornithine-coated glass slides and allowed to dry for at least 30 min at room temperature. Antiserum, prepared against mouse salivary gland NGF (electrophoretically pure), was diluted 1 in 4,000 in phosphate buffered saline (PBS) pH 7.4 containing 1% sheep serum and applied to the stretched tissues. Following 48 h incubation at 4 °C, the antiserum was removed and replaced by several changes of PBS. Fluorescein-labelled anti-rabbit IgG (Wellcome, Sydney) was used at a dilution of 1:100 and applied for 1 h then washed from the tissue before mounting. Scale bar, 10 μ m.

that a precise physiological role for the molecule cannot be ascribed until the cell types that produce it are known⁷⁻⁹. I report here the use of immunohistochemistry to localize endogenous NGF in the rat iris, a tissue in which there is sound biochemical evidence for the production of NGF activity¹⁰. Surprisingly, the results reveal that NGF can be detected readily in Schwann cells, but not in smooth muscle cells of the iris when it is sympathetically denervated or cultured.

Irides from adult albino rats (250–300 g body weight) of both sexes were used. After culture for four or six days in Eagles basal medium supplemented with 10% fetal calf serum, irides were removed, freeze-thawed and stretch-mounted on glass slides. Specific stain was clearly visible in all regions of the iris confined to discrete clusters of cells (Fig. 1). The distribution of these cells was similar to that of nerve fibres and bundles throughout the normal iris. Although many labelled cells were clustered in large bundles, individual fusiform cells were also

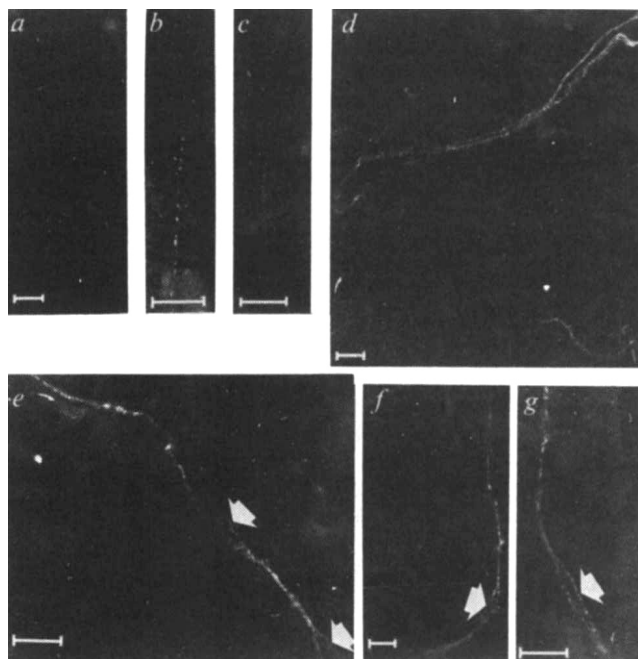


Fig. 2 Immunohistochemistry in the 6-day cultured rat iris. *a*, Absence of stain in tissue treated with pre-immune serum. *b, c*, Presence of stain in degenerating nerve fibres treated with dopamine β -hydroxylase antiserum. *d-g*, Presence of stain in strings of single cells stained with nerve growth factor antiserum. Arrows, nuclei discernable by the absence of stain. Staining procedure as for Fig. 1. Scale bar, 10 μ m.

present and were joined end-to-end in long strings with nuclei easily discernable by the absence of stain (Fig. 2). The resolution of the light microscope did not allow the exclusion of stain in degenerating nerve fibres. Even after four days of culture there were a few fragmented sympathetic nerve fibres present that could be visualized by immunohistochemistry with antiserum to dopamine β -hydroxylase¹¹ (Fig. 2). At no stage was anti-NGF stain visible in any smooth muscle fibres; irides that were processed without prior incubation gave negative results (Fig. 2). Cultured and non-cultured irides treated with pre-immune serum were also negative (Fig. 1). Identical results were seen in tissues fixed in picric acid solution prior to staining (Fig. 3). The specificity of the nerve growth factor antiserum was tested (Fig. 3); absorption of the antiserum with NGF-coupled beads blocked the immunohistochemical stain whereas absorption with insulin-coupled beads had no effect. Furthermore, the immunohistochemical stain could be demonstrated using dialysed antibody eluted from the NGF-coupled beads with 1 M acetic acid (Fig. 3).

Irides were next examined following removal of the ipsilateral superior cervical ganglion ten days previously. The effectiveness of sympathetic denervation was demonstrated in half of the iris by the absence of fluorescent nerve fibres following treatment with glyoxylic acid to induce catecholamine fluorescence¹². The remaining half of the iris was stained with NGF antiserum. The distribution of stain was similar to that seen after culture, although both the intensity of the stain and the number of positive cells was reduced considerably (Fig. 4). Contralateral irides treated with the antiserum were negative.

The results presented here provide the first histochemical evidence for the cellular localization of endogenous NGF outside the salivary gland. It is important therefore to examine the authenticity of the antisera used. The major antiserum was raised in rabbits against electrophoretically pure mouse salivary gland nerve growth factor prepared by the method of Mobley *et al.*¹³ (Fig. 4). The purified protein was active in the 1–10 ng ml⁻¹ range, tested by supporting dissociated chick sympathetic

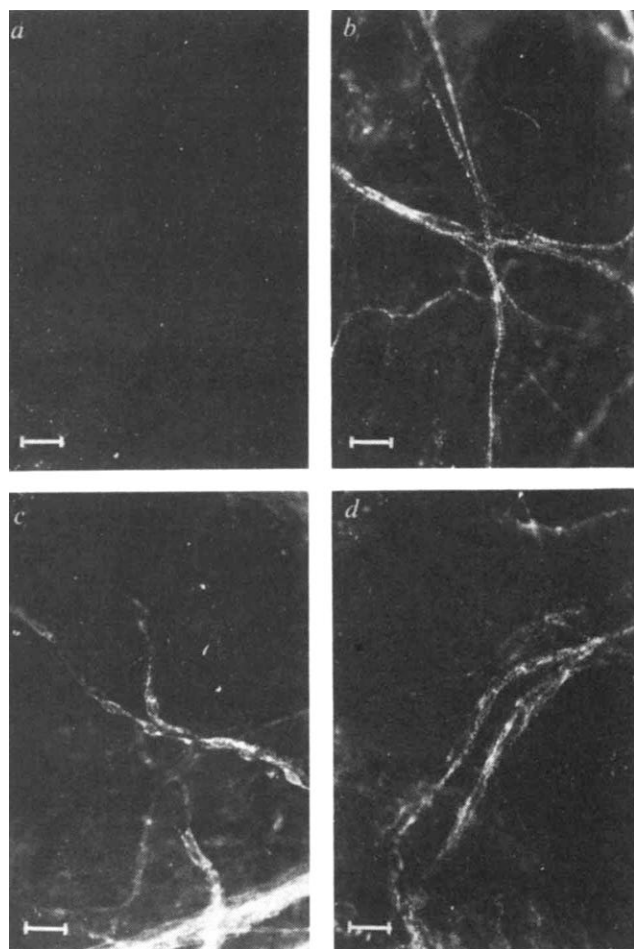


Fig. 3 Immunohistochemistry in the 6-day cultured rat iris. *a*, Absence of stain in an iris treated with antiserum previously absorbed with NGF-Sepharose; *b*, the presence of stain after treatment with whole antiserum; *c*, the presence of stain after treatment with antiserum absorbed with insulin-Sepharose; *d*, the presence of stain after treatment with antibody eluted from the NGF-coupled beads. Irises in *a*, *c* and *d* were fixed after freeze-thaw treatment and prior to staining in picric acid solution for 30 min (150 ml of saturated picric acid solution in 850 ml of PBS, pH 7.4); the iris in *b* was stained without prior fixation as described in Fig. 1.

Methods: Absorption of the antibody with solid phase antigen was performed as follows. Antiserum diluted to 1 in 4,000 was added to 100 μ l Sepharose 4B (Pharmacia) coupled to either NGF or insulin and incubated for 30 min at 37 °C in a shaking water bath. The supernatant was then removed from the beads and used immediately for immunohistochemistry. Purified antibody was eluted from the NGF-Sepharose with 1 M acetic acid and dialysed against PBS before use at a final protein concentration of 2.5 μ g ml⁻¹. Coupling of NGF and insulin to epoxy-activated Sepharose 4B was carried out according to the manufacturer's recommendations using 2 mg of each protein with 0.5 g of lyophilized gel. Scale bar, 10 μ m.

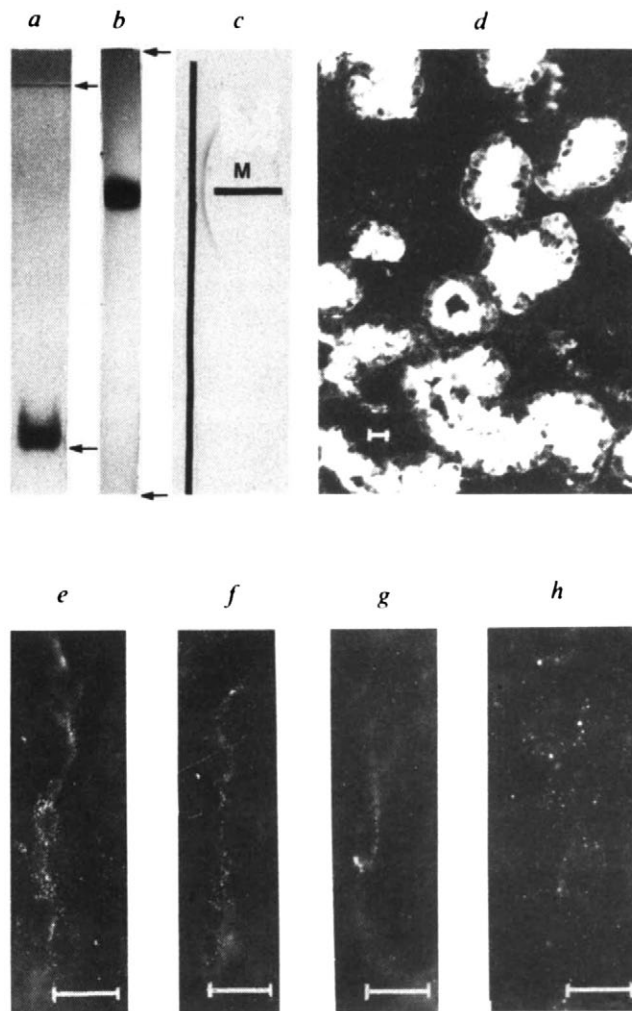


Fig. 4 *a-d*, Evidence for specificity of antiserum to NGF; *e-h*, NGF immunohistochemistry in the sympathetically denervated iris. *a*, Electrophoresis of purified mouse salivary gland nerve growth factor on 12.5% polyacrylamide discontinuous gels in SDS according to the method of Laemmli and Favre²⁰ with a protein load of 30 μ g. *b*, 20 μ g purified NGF on a 30% polyacrylamide gel, pH 10.3, according to the method of Smith, Varon and Shooter²¹. Arrows, position of the origins and dye fronts. *c*, Immunoelectrophoresis against crude mouse salivary gland extract. Electrophoresis was carried out on Corning Universal electrophoretic agarose film (Corning) in 0.065 M barbital buffer, pH 8.6 for 45 min at 200 V with constant cooling. Under these conditions, mouse salivary gland 2.5 S NGF does not migrate significantly from the origin. The film was stained after washing with Coomassie blue. *d*, Immunohistochemical localization of NGF in submaxillary gland of male mouse. Anaesthetized male mice were perfused with a solution of 4% paraformaldehyde in 0.1 M phosphate buffer, pH 7.4, the submaxillary glands removed, cut into small blocks and immersed for 3-4 h in the same fixative. The tissue blocks were then rinsed overnight in 5% sucrose in 0.1 M phosphate buffer pH 7.4 and sectioned at 10 μ m on a cryostat. Sections were stained for 48 h at 4 °C with anti-NGF serum diluted to 1 in 4,000 as described in Fig. 1. The fluorescent stain was most intense in apical regions of granular duct cells and absent in acinar cells. *e-h*, fluorescent cells stained with anti-NGF in irides sympathetically denervated 10 days previously. Scale bar, 10 μ m.

neurones in culture, elevating ornithine decarboxylase activity in newborn rat superior cervical ganglia and in increasing the number of surviving neurones in sympathetic ganglia when administered to developing chick embryos. The resulting antiserum was monospecific against NGF, shown by immunoelectrophoresis against both the purified factor and crude homogenates of male mouse salivary glands; it also stains granular duct cells in this tissue (Fig. 4). This antiserum blocks the survival of sympathetic neurones cultured in the presence of NGF¹⁴. Antiserum and commercial antibodies (Laref, Switzerland), purified by chromatography on an NGF affinity column, gave identical results whereas normal sera from several different rabbits gave negative results. Thus the staining in the figures is

probably a result of the interaction of NGF antibodies with an endogenous NGF-like molecule.

The staining present in denervated or cultured irides is localized to a small proportion of the total cells present that are confined to nerve fibre tracts. I suggest here that these are Schwann cells and possibly associated nerve fibres. In the absence of electron microscope evidence, there are several rea-

sons for this conclusion. First, the stained cells clearly are not smooth muscle fibres as their morphology and distribution are distinct from these cells. Second, their morphology is identical to Schwann cells and they are arranged either in clusters or as elongated single cells in bundles resembling nerve fibre bundles. Third, although fibroblasts are present in nerve bundles, their total number is only about 5% of the number of Schwann cells¹⁵. It is unlikely that fibroblasts would be confined to nerve bundles and yet no isolated stained cells were observed outside the bundles. However, this does not exclude the possibility that some positive cells are fibroblasts. Finally, the production of NGF by purified cultures of Schwann cells^{16,17} and isolated peripheral nerve segments¹⁸ is established clearly.

Although it is clear that the anti-NGF stain is associated with nerve tracts throughout both the radial and sphincter muscles, it was not possible to establish which type of nerves are involved. Specific stain was present in both fine nerve tracts and thick bundles; the density of the fibre tracts stained throughout the iris was only slightly less than the density of sympathetic fibres in a non-cultured iris revealed by glyoxylic acid-induced catecholamine fluorescence. In these cultured irides the nerve axons and terminals detached from their cell bodies are degenerating; immunohistochemical localization with dopamine β -hydroxylase antiserum has confirmed that many of the sympathetic nerve fibres have been lost. However, experiments are currently in progress to double-label tissues with NGF and substance P, dopamine β -hydroxylase or choline acetyltransferase antisera to discover whether sensory, sympathetic or parasympathetic fibres are involved.

The production of NGF by Schwann cells in the denervated iris suggests that damaged nerves receive trophic support in the old fibre tracts during regeneration. It also suggests that after injury the nerve returns to an embryonic-like sensitivity. Certainly, regenerating sympathetic nerves are sensitive to NGF⁴, whereas intact nerves in the mature animal show little response to the factor⁷.

Many target tissues of the sympathetic and sensory nervous systems are known to produce NGF in culture¹⁹. This observation has been used to strengthen the proposal that neuronal death occurs during development as a result of competition between growing axons for NGF secreted from target cells, such as smooth and cardiac muscle fibres. However, extrapolation of the present findings to developing nerves would suggest that the nerves are supported during the period of axonal outgrowth by NGF from Schwann cells and that at the time of innervation, the target cells provide a second trophic molecule that is not NGF.

I thank Miss Fiona Renton for technical assistance and Dr Ian Hendry for supplying a sample of antiserum to nerve growth factor. The work was supported by the NH and MRC of Australia and by equipment grants from the MDA of South Australia and the Flinders Medical Centre Research Foundation.

Received 17 August 1983; accepted 11 September 1984.

1. Levi-Montalcini, R. & Angeletti, P. U. *Physiol. Rev.* **48**, 534-569 (1968).
2. Hendry, I. A. in *Reviews of Neuroscience* Vol. 2 (eds Ehrenpreis, S. & Kopin, I. J.) 149-194 (Raven, New York, 1976).
3. Levi-Montalcini, R. & Booker, B. *Proc. natn. Acad. Sci. U.S.A.* **46**, 373-383 (1960).
4. Hendry, I. A. & Campbell, J. *J. Neurocytol.* **5**, 351-360 (1976).
5. Levi-Montalcini, R. & Booker, B. *Proc. natn. Acad. Sci. U.S.A.* **46**, 384-391 (1960).
6. Levi-Montalcini, R. & Angeletti, P. U. *Pharm. Rev.* **18**, 619-628 (1966).
7. Harper, G. P. & Thoenen, H. *J. Neurochem.* **34**, 5-16 (1980).
8. Hamburger, V., Brunso-Bectold, J. K. & Yip, J. W. *J. Neurosci.* **1**, 60-71 (1981).
9. Bradshaw, R. A. *Rev. Biochem.* **41**, 91-216 (1978).
10. Ebendal, T., Olsen, L., Seiger, A. & Hedlund, K.-O. *Nature* **286**, 25-28 (1980).
11. Rush, R. A. in *Methods in Neurosciences* (ed. Cuelllo, A. C.) 139-158 (Wiley, Chichester 1983).
12. Axelsson, S., Bjorklund, A., Falck, B., Lindvall, O. & Svensson, L. A. *Acta physiol. scand.* **87**, 57-62 (1973).
13. Mobley, W. C., Schenker, A. & Shooter, E. M. *Biochemistry* **15**, 5543-5551 (1976).
14. Rush, R. A., Abrahamson, I. K., Belford, D. A., Murdoch, S. Y. & Wilson, P. A. in preparation.
15. Causey, G. & Barton, A. A. *Brain* **82**, 594-598 (1959).
16. Burnham, P., Raiborn, C. & Baron, S. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3556-3560 (1972).
17. Varon, S. S. & Bunge, R. P. A. *Rev. Neurosci.* **1**, 327-361 (1978).
18. Richardson, P. M. & Ebendal, T. *Brain Res.* **246**, 57-64 (1982).
19. Harper, G. P., Pearce, F. L. & Vernon, C. *Dev. Biol.* **77**, 391-402 (1980).
20. Laemmli, U. K. & Favre, M. *J. molec. Biol.* **90**, 575-599 (1983).
21. Smith, A. P., Varon, S. & Shooter, E. M. *Biochemistry* **7**, 3259-3268 (1968).

Aetiology of AIDS—antibodies to human T-cell leukaemia virus (type III) in haemophiliacs

L. W. Kitchen*, F. Barin*, J. L. Sullivan†, M. F. McLane*, D. B. Brettler†, P. H. Levine† & M. Essex*

* Department of Cancer Biology, Harvard School of Public Health, 665 Huntington Avenue, Boston, Massachusetts 02115, USA

† Departments of Pediatrics and Medicine, University of Massachusetts Medical School, New England Comprehensive Hemophiliac Center, and the Department of Medicine, Worcester Memorial Hospital, Worcester, Massachusetts 01605, USA

Human T-cell leukaemia virus type III (HTLV-III) is suspected of having a key role in the pathogenesis of acquired immune deficiency syndrome (AIDS)¹⁻⁴. Epidemiological data suggest that AIDS is transmitted by an infectious agent through intimate contact with body secretions, blood or blood products⁵. To maintain haemostasis, many haemophiliac patients depend on commercially prepared clotting concentrates made from large multi-donor plasma pools and are thus at increased risk of developing the disease^{6,7}. We report here that, using indirect membrane immunofluorescence and radioimmunoprecipitation with SDS-polyacrylamide gel electrophoresis, we have detected antibodies to HTLV-III in 30 of 47 (64%) asymptomatic haemophiliacs and in all of three haemophiliacs who either had or soon developed AIDS. Of 34 samples drawn before 1984, 18 (53%) were antibody-positive, whereas of 16 samples drawn during 1984, 15 (94%) were positive ($P \leq 0.002$). These data suggest that exposure to HTLV-III antigen is widespread among asymptomatic haemophiliacs.

We tested samples from 50 haemophiliacs from the New England Area Comprehensive Hemophilia Center in Worcester, Massachusetts; 44 of the patients had haemophilia A and 6 had haemophilia B. All haemophilia A patients had moderate or severe disease and received home infusions of commercial factor VIII concentrate. The samples were collected during 1983-84 from individuals ranging in age from 1 to 69 yr. All were asymptomatic at the time of serum collection, except for a single patient who had clinical AIDS. Two more individuals developed AIDS within 3-6 months (confirmed by the Center for Disease Control).

All sera were initially examined at a 1:4 dilution by indirect membrane immunofluorescence (IMI) on the HTLV-III-infected line, designated H9/HTLV-III, and the uninfected line, designated H9. The method used was that described earlier, except that the cells were inactivated with 2% paraformaldehyde in phosphate-buffered saline as a biohazard precaution before the tests were scored. All 50 samples were subsequently tested using radioimmunoprecipitation with SDS-polyacrylamide gel electrophoresis (RIP-SDS/PAGE). Both uninfected H9 and HTLV-III-infected cells were metabolically labelled for 8 h with ³⁵S-cysteine (100 μ Ci ml⁻¹) according to details described elsewhere^{8,9}.

Of 50 serum samples tested, 33 (66%) reacted with the HTLV-III-infected cells by IMI, but not with the uninfected cells (Table 1, Fig. 1). The same 33 sera showed immunoprecipitation of proteins that migrate in the same position as those regularly detected with reference antiserum to HTLV-III core proteins (provided by Dr R. C. Gallo). Samples from 29 healthy laboratory workers were all negative when tested by both membrane immunofluorescence and immunoprecipitation. The antigens most consistently detected by the 33 positive sera included proteins of molecular weights 24,000 (p24), 41,000 (gp41), 53,000 (p53), 120,000 (gp120) and 160,000 (gp160); these are the proteins that Gallo and colleagues¹⁻⁴ and workers in our laboratory (T. H. Lee and M.E., manuscript in preparation) have found are most frequently detected by human HTLV-III antibodies from AIDS patients. Immunoprecipitation of the gp41, gp120 and gp160 proteins occurred when using a lentil

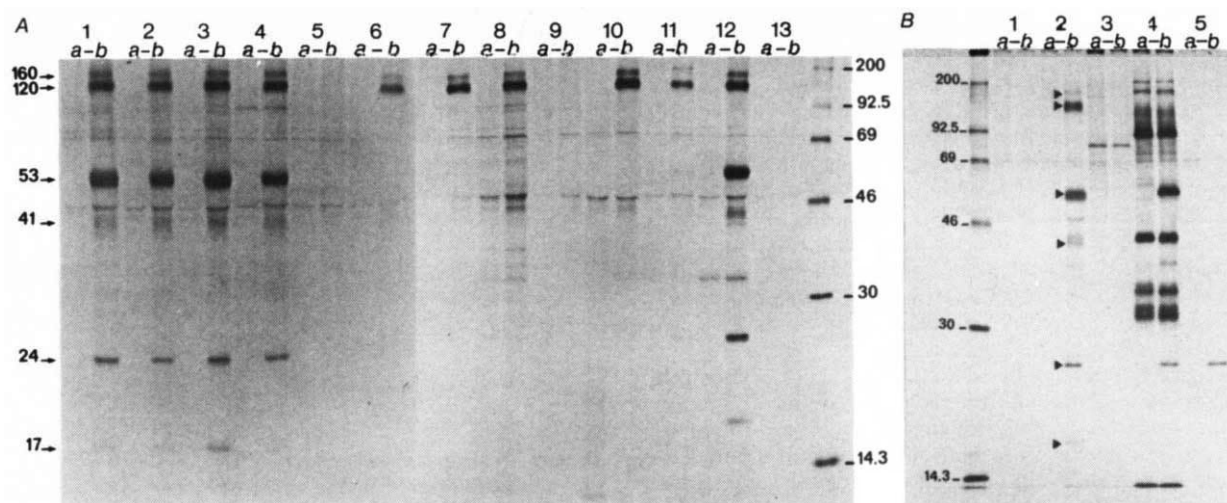


Fig. 1 A, 35 S-cysteine-labelled antigens precipitated from HTLV-III-negative H9 cells (a) and HTLV-III-infected H9 cells (b) by serum samples from asymptomatic haemophiliacs (lanes 1-5, 9-11). Lane 12: a positive control AIDS patient. Lane 13: a negative control (healthy laboratory worker). Lanes 6 and 7: asymptomatic haemophiliacs who developed AIDS within 3-6 months. Lane 8: a haemophiliac AIDS patient. Numbers on the left refer to proteins discussed in the text; numbers on the right refer to molecular weight markers ($\times 10^{-3}$). B, 35 S-cysteine-labelled antigens precipitated from HTLV-III-negative cells (a) and HTLV-III-infected H9 cells (b). Serum samples were from: lane 1, a negative control (same as lane 13 in A); lane 2, a positive control AIDS patient (same as lane 12 in A); lane 3, a normal rabbit; lane 4, a reference rabbit antiserum to disrupted HTLV-III³; lane 5, a monoclonal antibody to HTLV-III p24. Numbers on the left refer to molecular weight markers ($\times 10^{-3}$).

lectin-bound, glycoprotein-enriched cell lysate fraction as described earlier, which suggests that these antigens are glycoprotein in nature¹⁰. Seventeen of the serum samples failed to react by either IMI or RIP-SDS/PAGE. Of the 33 serum samples that were positive for antibody to HTLV-III-related glycoproteins, 23 also proved positive for antibody to virus core antigens. p24 is the major virion core protein and p53 is presumed to be the gag gene-encoded polyprotein because, as Gallo and his co-workers¹⁻⁴ and workers in our laboratory (T. H. Lee and M.E., manuscript in preparation) have found, it is readily precipitated with reference antiserum to disrupted HTLV-III. Some haemophiliac samples precipitated a protein of 17,000 molecular weight from H9/HTLV-III cells; this protein is also precipitated by a reference rabbit antiserum to HTLV-III core proteins.

The 50 haemophiliacs tested included three who developed AIDS. All three tested positive by both procedures. Patients who tested positive ranged in age from 4 to 57 yr and included patients with haemophilia A (32 of 44) and haemophilia B (1 of 6). When divided according to the time of sample collection, 18 of 34 (53%) of individuals samples before 1984 were positive whereas 15 of 16 (94%) of those studied during 1984 were positive (one-sided P value ≤ 0.002).

Evidence has been accumulating which suggests that a member of the human T-cell leukaemia retrovirus (HTLV) family is involved in the pathogenesis of AIDS^{1-5,11-17}. Other retroviruses have been linked with naturally occurring immunosuppression and leukaemia in cats¹⁸, and a virus closely related to HTLV

has been associated with leukaemias and lymphoproliferative diseases that have occurred as part of an AIDS-type outbreak in macaque monkeys⁸.

We have previously performed IMI on serum samples of haemophiliacs to test for the presence of antibodies to membrane antigens of cells infected with HTLV-I (HTLV-MA)^{9,11,19}. Haemophiliacs were also tested for antibodies to cytomegalovirus (CMV) and Epstein-Barr virus (EBV)²⁰. Sixty-nine per cent of haemophiliacs had antibodies to EBV and 42% had antibodies to CMV. Approximately 12% of 172 asymptomatic haemophiliacs from four cities tested positive for HTLV-MA; specificity was confirmed with RIP-SDS/PAGE⁹. Because a significant proportion of the haemophiliacs from New York had evidence of antibodies to HTLV-MA before 1980, when AIDS was first described in the United States, it seemed likely that some had been exposed to a conventional strain of HTLV rather than to HTLV-III, the strain that seems to be the aetiological agent of AIDS²¹.

The present results suggest that exposure to HTLV-III is widespread in asymptomatic haemophiliacs. Earlier studies indicated that antibodies to HTLV-I and HTLV-II envelope gene glycoproteins are associated with the development of AIDS^{11,15,22}, presumably because the antibodies are at least partially cross-reactive with HTLV-III proteins¹⁻⁴. Both HTLV-I and AIDS, which is thought to be caused by HTLV-III, seem to be transmitted by blood transfusions and blood products^{15,23-26}. In a previous study, two of three haemophiliacs with AIDS had antibodies to HTLV-MA of type I⁹ and in the

Table 1 Presence of antibodies to HTLV-III in haemophiliacs and healthy laboratory workers

Category	No. tested	HTLV-MA	gp41/gp120/gp160	p24/p53
Haemophilia A	44	32 (73)	32 (73)	23 (52)
Haemophilia B	6	1 (17)	1 (17)	0 (0)
1984	16	15 (94)	15 (94)	13 (81)
Pre-1984	34	18 (53)	18 (53)	10 (29)
Asymptomatic	47	30 (64)	30 (64)	23 (49)
Haemophilia with AIDS	3	3 (100)	3 (100)	0 (0)
Total haemophiliacs	50	33 (66)	33 (66)	23 (46)
Healthy laboratory workers	29	0 (0)	0 (0)	0 (0)

Antibodies to HTLV-MA were determined by IMI; those to gp41/gp120/gp160 and p24/p53 were determined by RIP-SDS/PAGE.

present study two asymptomatic haemophiliacs who subsequently developed AIDS had antibodies to HTLV-III.

The observation that the proportion of asymptomatic haemophiliacs with antibodies to HTLV-III may have increased from 1983 to 1984 deserves further study using sequential samples from the same individuals. As the incubation period for the development of AIDS following blood transfusion has been estimated to range from 10 months to 4 yr²⁶, further studies are needed to evaluate the impact of this observation. Antibody positivity seems to be an indicator of viral infection in most healthy and leukaemic individuals who were naturally exposed to HTLV-I²⁷⁻²⁹ as well as in AIDS patients and healthy male homosexuals exposed to HTLV-III¹⁻⁴. However, it is possible that a significant proportion of asymptomatic haemophiliacs might be exposed only to inactivated HTLV-III rather than to live virus, owing to the manufacturing process involved in the preparation of commercial factor VIII concentrate. In addition,

it is clear that those individuals exposed to factor IX preparations are also at risk for the development of antibodies to HTLV-III. Further studies are needed to determine whether the presence of antibodies to HTLV-III indicates immunization, infection or an asymptomatic carrier state. In addition, the relationship between HTLV-III seropositivity and immunological dysfunction, which has been shown to be widespread among apparently healthy haemophiliacs^{30,31}, requires investigation.

We thank Dr R. C. Gallo for the reference cell cultures and the reference antisera. This research was supported by National Research Service Award ST 32CA9382 (L.W.K.), Association pour le Développement de la Recherche sur le Cancer, Fondation pour la Recherche Médicale, and Philippe Foundation (F.B.), NIH grants AI21161 (J.L.S.) and CA37466 and CA18216 (M.E.). J.L.S. is an established investigator of the American Heart Association.

Received 30 August; accepted 26 September 1984.

1. Popovic, M., Sarngadharan, M. G., Read, E. & Gallo, R. C. *Science* **224**, 497-500 (1984).
2. Gallo, R. C. *et al. Science* **224**, 500-503 (1984).
3. Schupbach, J. *et al. Science* **224**, 503-505 (1984).
4. Sarngadharan, M. G., Popovic, M., Bruch, L., Schupbach, J. & Gallo, R. C. *Science* **224**, 506-508 (1984).
5. Francis, D. P., Curran, J. & Essex, M. *J. int. Cancer Inst.* **71**, 1-4 (1983).
6. Davis, K. C., Horsburgh, C. R., Jr, Hasibu, U., Schocket, A. L. & Kirkpatrick, C. H. *Ann. intern. Med.* **98**, 284-286 (1983).
7. Elliott, J. L. *et al. Ann. intern. Med.* **98**, 290-293 (1983).
8. Homma, T. *et al. Science* **225**, 716-718 (1984).
9. Essex, M. *et al. Science* **221**, 1061-1064 (1983).
10. Lee, T. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3856-3860 (1984).
11. Essex, M. *et al. Science* **220**, 859-862 (1983).
12. Gelmann, E. P. *et al. Science* **220**, 862-865 (1983).
13. Gallo, R. C. *et al. Science* **220**, 865-867 (1983).
14. Barre-Sinoussi, F. *et al. Science* **220**, 868-871 (1983).
15. Jaffe, H. W. *et al. Science* **223**, 1309-1312 (1984).

16. Feorino, P. M. *et al. Science* **225**, 69-72 (1984).
17. Kalyanaram, V. S. *et al. Science* **225**, 321-323 (1984).
18. Essex, M., Hardy, W. D., Jr, Cotter, S. M., Jakowski, R. M. & Sliski, A. *Infect. Immun.* **11**, 470-475 (1975).
19. Kreiss, J. K. *et al. Ann. intern. Med.* **100**, 178-182 (1984).
20. Cheeseman, S. H., Sullivan, J. L., Brettler, D. B. & Levine, P. H. *J. Am. med. Ass.* **252**, 83-85 (1984).
21. Evatt, B. L. *et al. Lancet* **ii**, 698-701 (1983).
22. Lee, T. H. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
23. Miyoshi, I. *et al. Lancet* **i**, 683-684 (1982).
24. Saxinger, W. C. & Gallo, R. C. *Lancet* **i**, 1074 (1982).
25. Okochi, K., Sato, H. & Hinuma, Y. *Vox Sang* **46**, 245-253 (1984).
26. Curran, J. W. *et al. New Engl. J. Med.* **310**, 69-75 (1984).
27. Kalyanaram, V. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 1653-1657 (1982).
28. Gotoh, Y. I., Sugamura, K. & Hinuma, Y. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4780-4782 (1982).
29. Robert-Guroff, M. *et al. Science* **215**, 975-978 (1982).
30. Lederman, M. M., Ratnoff, O. D., Scillian, J. J., Jones, P. K. & Schacter, B. *New Eng. J. Med.* **308**, 79-82 (1983).
31. Menitove, J. E. *et al. New Engl. J. Med.* **308**, 83-86 (1983).

Significance of herpesvirus immediate early gene expression in cellular immunity to cytomegalovirus infection

Matthias J. Reddehase & Ulrich H. Koszinowski

Federal Research Centre for Virus Diseases of Animals, Tübingen, PO Box 1149, D-7400 Tübingen 1, FRG

Interstitial pneumonia linked with reactivation of latent human cytomegalovirus due to iatrogenic immunosuppression can be a serious complication of bone marrow transplantation therapy of aplastic anaemia and acute leukaemia¹. Cellular immunity plays a critical role in the immune surveillance of inapparent cytomegalovirus infections in man and the mouse¹⁻⁷. The molecular basis of latency, however, and the interaction between latently or recurrently infected cells and the immune system of the host are poorly understood. We have detected a so far unknown antigen in the mouse model. This antigen is found in infected cells in association with the expression of the herpesvirus 'immediate early' genes and is recognized by cytolytic T lymphocytes (CTL)⁸. We now demonstrate that an unexpectedly high proportion of the CTL precursors generated *in vivo* during acute murine cytomegalovirus infection are specific for cells that selectively synthesize immediate early proteins, indicating an immunodominant role of viral non-structural proteins.

According to the static state hypothesis of latency⁹, it is assumed that viral progeny are not produced because the transcription process of the virus is interrupted reversibly at an early stage in the cycle of virus multiplication. In herpes virus infections of permissive cells *in vitro*, analysed in detail for herpes simplex viruses^{10,11}, viral protein synthesis is coordinately regulated and starts with the expression of the 'immediate early' (IE) genes, which are transcribed by host RNA polymerase II. The IE proteins exert important control functions as exemplified by their ability to act as transcriptional activators^{12,13} and they need to be present throughout the replication cycle of herpes-

Table 1 Determination of the IL-CTLp frequency by LDCC

Dose of lectin ($\mu\text{g ml}^{-1}$)	IL-CTLp frequency	No. of IL-CTLp per 10^6 lymphocytes (95% confidence limits)	pMC*	
Con A	5	1/5,900	171 (126-215)	0.23
	10	1/2,800	359 (217-502)	0.11
	20	1/2,600	383 (270-495)	0.50
PHA	5	1/5,600	179 (124-234)	0.24
	10	1/9,500	105 (71-140)	0.52
	20	1/7,600	132 (93-171)	0.24

BALB/c (*H-2^d* haplotype) mice were infected at the left hind footpad with 10^5 PFU of tissue-culture propagated MCMV (ATCC VR-194 Smith strain) and 8 days later lymphocytes derived from the draining popliteal lymph nodes were tested for the frequency of activated CTLp in a limiting dilution assay²⁰. In brief, graded numbers of lymphocytes were cultivated for 6 days in 96-well round-bottomed microtitre plates with no addition of viral antigen using medium MEM- α (Gibco) supplemented with 10% vol. fetal calf serum and 20% vol. interleukin-2 containing supernatant. Starting the dilution series with 1×10^5 lymphocytes, 7 log₂ dilution steps were performed with 24 replicate microcultures each. To detect all CTL clones irrespective of their antigen-specificity, the LDCC assay was applied. Various doses of the lectins Con A (Pharmacia) or PHA (Flow Laboratories) were added to 10^3 ⁵¹Cr-labelled, noninfected MEF and CTL at the beginning of a standard 3-h ⁵¹Cr release assay. Precursor frequencies were estimated from the proportion of nonresponding microcultures using the minimum χ^2 (MC)³² and maximum likelihood (ML)³³ methods. Only the results of the MC evaluation are shown.

* Probability calculated by MC method. According to the proposition of Taswell³² an estimate was accepted for a probability of $P > 0.05$ and a ML-MC deviation of $< 10\%$.

viruses to permit transcription of the 'early' and 'late' genes¹⁴. Interference with viral transcription at the stage of the expression of IE genes may be involved in establishing viral latency; hence many efforts have been made to clone and map the IE regions of herpesvirus genomes, including those of human^{15,16} and murine cytomegalovirus (MCMV)¹⁷⁻¹⁹. In particular, re-expression of IE genes must be one of the first events in virus reactivation from latency. Selective synthesis of IE proteins leads

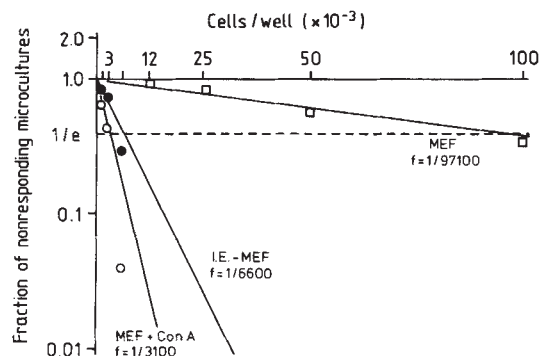


Fig. 1 High frequency of LYDIEA-specific IL-CTLp. CTL were generated by *in vitro* interleukin-mediated expansion of *in vivo* activated interleukin-receptive CTLp (IL-CTLp) as outlined in Table 1. Second-passage mouse embryo fibroblasts (MEF) were used as a source of target cells for the ^{51}Cr -release assay. To determine the negative or positive reference frequencies, noninfected MEF were used either in the absence of lectin or with addition of $20\text{ }\mu\text{g ml}^{-1}$ Con A, respectively. Infected target cells after selective expression of immediate early genes (IE-MEF) served to determine the frequency of IL-CTLp specific for the immediate early antigen (LYDIEA) of MCMV. Details of the limiting dilution assay are given in Table 1 and details of the target cell preparation in Table 2.

to the expression of an antigen on the cell surface⁸. This antigen is recognized by CTL and was therefore designated lymphocyte-detected IE antigen (LYDIEA). LYDIEA-specific CTL represented a distinct subset, separate from those CTL recognizing viral structural antigens⁸. The mere fact, however, that during MCMV infection LYDIEA-specific CTL are generated did not allow definite conclusions about the relative significance of the immune response to LYDIEA in comparison with the response to other antigens specified by MCMV.

The solution of two experimental problems was a prerequisite for attempts to evaluate the significance of LYDIEA. First, *in vivo* primed CTL precursors (CTLp) can be detected only after clonal expansion *in vitro* by testing their cytolytically active progeny. This expansion has to be performed in the absence of antigen because restimulation *in vitro* may favour or disfavour the growth of CTLp specific for certain antigens and thereby cause a distortion of the numerical *in vivo* representation of specificities. The expansion of interleukin-receptive CTLp (IL-

CTLp) by interleukin solves this problem as that procedure selects cells only on the basis of differentiation²⁰.

The second problem arises from the fact that progeny CTL are usually detected by their ability to recognize target cells carrying the relevant antigens. This recognition requires sufficient amounts of antigen to be present at the cell membrane²¹. Selective and sufficient expression of LYDIEA can be achieved⁸, but it is not certain in a particular experiment whether each of the other viral antigens is also expressed sufficiently. Therefore, to detect all IL-CTLp generated during MCMV infection regardless of their antigen-specificity, the lectin-dependent cellular cytotoxicity (LDCC) assay was applied. In this assay the T-cell mitogenic lectins concanavalin A (Con A) and phytohaemagglutinin (PHA) are used to bypass the requirement for specific interaction between CTL and target cells^{22,23}. The highest frequency of IL-CTLp was measured when ConA was added at doses of $10\text{--}20\text{ }\mu\text{g ml}^{-1}$ (Table 1). The fact that the corresponding frequency in unprimed mice ranged from 1 in 70,400 to 1 in 91,400 lymphocytes (data not shown) proved that the CTL detected in the lectin-dependent cellular cytotoxicity (LDCC) assay were derived from *in vivo* sensitized precursors. This condition was then used to assess the total frequency of IL-CTLp.

The results of a representative limiting dilution assay are shown in Fig. 1 and those of an independent second experiment summarized in Table 2. The frequency of LYDIEA-specific IL-CTLp present in a draining popliteal lymph node was found to be 1 in 6,600 lymphocytes (Fig. 1) or 1 in 6,900 lymphocytes, respectively (Table 2). This determination was carried out on target cells that selectively expressed high amounts of IE proteins. The selective transcription of IE genes in target cells was achieved by infection in the presence of the translation inhibitor cycloheximide that led to an accumulation of IE RNA. The subsequent release from cycloheximide and further cultivation in the presence of the transcription inhibitor actinomycin D then resulted in an enhanced synthesis of IE proteins and simultaneously prevented further viral transcription^{8,19}.

During noninterrupted (productive) infection, the synthesis of IE proteins is followed by the synthesis of so-called early proteins and, after viral DNA replication, of late proteins which include the major viral structural proteins^{10,24}. Among those especially the glycoproteins of the virion envelope should contribute significantly to the antigenicity of infected cells. Nevertheless, when tested on cells corresponding to the late

Table 2 IL-CTLp frequencies as determined by antigen-specific or lectin-dependent cellular cytotoxicity

Target cell (virus dose; inhibitors)	IL-CTLp frequency	No. of IL-CTLp per 10^6 lymphocytes (95% confidence limits)	pMC
MEF (—; CH-Act.D)	1/166,400	6 (4–8)	0.35
MEF + Con A (—; CH-Act.D)	1/2,700	371 (226–516)	0.27
IE-MEF (0.05; CH-Act.D)	1/6,900	146 (97–195)	0.16
Late-MEF (0.05; -Act.D)	1/11,300	89 (60–118)	0.24
UV-MCMV-MEF (0.05; CH-Act.D)	1/85,300	12 (8–16)	0.35
UV-MCMV-MEF (0.5; CH-Act.D)	1/44,300	23 (12–33)	0.16
UV-MCMV-MEF (5; CH-Act.D)	1/14,900	67 (48–86)	0.08

Second-passage MEF were used as target cells throughout. Infection was performed under centrifugal enhancement conditions³⁴ ($800g$ for 30 min) with the indicated virus dose of infection that is given in plaque-forming units (PFU) per cell as determined in a conventional plaque titration assay. The centrifugation enhances the efficiency of infection by 20- to 80-fold and thus a virus dose of 0.05 PFU corresponds with an effective multiplicity of about 1–4 PFU per cell. At a multiplicity of 100–400 PFU per cell target cells started to disintegrate and therefore the application of higher doses was precluded. For the preparation of IE targets and the appropriate controls, infection and 3 h of cultivation were performed in the presence of cycloheximide (CH, $50\text{ }\mu\text{g ml}^{-1}$). Thereafter, CH was washed out and actinomycin D (Act.D, $5\text{ }\mu\text{g ml}^{-1}$) was added. Only for the preparation of the late-MEF target, infection was carried out in the absence of CH and viral replication was allowed to proceed for 20 h until transcription was stopped by addition of actinomycin D. To test lectin-dependent cytotoxicity, Con A was added during the effector phase at a final dose of $20\text{ }\mu\text{g/ml}$. For mock-infection of MEF, partially purified stock virus (2×10^8 PFU ml^{-1})²⁰ was diluted with phosphate-buffered saline to a final concentration of 1×10^7 PFU ml^{-1} and then exposed as an 1-mm layer to a 10-min irradiation with a 90 W UV-illuminator at 254 nm. The distance between specimen surface and ultraviolet-screen was adjusted to give an intensity of $30\text{ J m}^{-2}\text{ s}^{-1}$. To control the effect of irradiation on viral protein synthesis³⁵, MEF were infected with the highest dose of virus used in the experiments (5 PFU equivalents per cell) and IE protein synthesis was enhanced by inhibitor treatment. Monoclonal antibody 6/58/1 (to be described elsewhere) was then used to detect IE nuclear antigen by indirect immunofluorescence. Irradiation of virus for 2.5 min ($4,500\text{ J m}^{-2}$) reduced the number of IE antigen positive cells to less than 1% and after 5 min of irradiation viral protein synthesis was abolished in all cells.

stage of viral replication, the frequency of IL-CTLp was only 1 in 12,800 lymphocytes (not depicted in Fig. 1) or 1 in 11,300 lymphocytes, respectively (Table 2).

In addition to *de novo* synthesized antigens, virion proteins or glycoproteins inserted into the cell membrane during the process of penetration can directly account for an antigenicity of MCMV-infected cells. This has been demonstrated previously by infection of cells in the continuous presence of actinomycin D or cycloheximide⁸. Different treatment, mock-infection with UV-inactivated virus, still permits the synthesis of cellular proteins which might be induced by viral structural proteins and could also contribute to target formation. This control is therefore indicated to discriminate between target antigens induced by the IE activity of the viral genome and those antigens whose expression does not require IE protein synthesis. When mock-infection of targets was performed with a dose of ultraviolet-inactivated virus equivalent to the dose of infectious virus applied for preparing the IE target, the frequency was only 1 IL-CTLp in 85,300 lymphocytes (Table 2). This result proved that the high frequency of 1 IL-CTLp in 6,900 lymphocytes, detected with the IE target, was indeed due to the expression of an IE antigen and thus referred to LYDIEA-specific IL-CTLp. This conclusion is further confirmed by the finding that in the absence of viral genome-induced protein synthesis even a hundred-fold increase in the virus dose did not allow to detect the number of IL-CTLp that could be detected with the IE target or with the productively infected target (Table 2).

The total frequency of IL-CTLp in these experiments was found to be 1 in 3,100 lymphocytes (Fig. 1) or 1 in 2,700 lymphocytes, respectively (Table 2). Hence, based on the determination of the total IL-CTLp frequency, the LYDIEA-specific subset comprised 40–50% of all IL-CTLp generated during infection.

The finding that such a substantial proportion of the IL-CTLp recognize an antigen expressed during the IE phase of viral replication was unexpected because transcription at that stage is restricted to only few genes, while extensive transcription from the whole genome occurs at later stages¹⁹.

Three possibilities might explain the high frequency of LYDIEA-specific IL-CTLp: First, the primary T cell specificity repertoire of BALB/c mice may already contain a high relative number of precursors specific for this antigen. There is no approach to test this possibility as limiting dilution experiments have revealed an extremely low frequency of primary CTLp specific for MCMV²⁰ or herpes simplex virus²⁵ and that is why a reliable comparison between particular antigens is not feasible.

Second, the expression of antigens may be different in the tissues of the infected animal than predicted from the productive infection of embryonic fibroblasts *in vitro*. Since in the immunocompetent host MCMV does not replicate at the site of intraplantar infection, whereas after immunosuppression fibroblasts of the footpad tissue are productively infected (manuscript in preparation), the immune system is involved in terminating the *in vivo* viral replication at an early stage. This could imply a preferential sensitization of LYDIEA-specific CTLp and explain the high proportion of LYDIEA-specific IL-CTLp in the draining lymph node. We are now testing the amount of IE gene expression by *in situ* hybridization using cloned viral DNA fragments¹⁹.

Third, the IE antigen may be highly immunogenic. Its molecular identity has first to be defined in order to evaluate this alternative. At present it is known that the expression of LYDIEA requires transcription and translation⁸ and coincides with the synthesis of a phosphorylated but not glycosylated protein of 89,000 apparent relative molecular mass. This protein and its post-translational modification products are the dominant IE proteins located in the nucleus and cytoplasm of the infected cell (G. M. Keil *et al.*, in preparation). Since recognition by CTL must be considered as a *prima facie* evidence for surface localization of an antigen, the actual relationship between these serologically detectable intracellular proteins and the antigen LYDIEA requires further analysis. The presence of

nonglycosylated viral proteins, located predominantly in the nucleus and cytoplasm, coincides with the recognition of infected cells by CTL also in other virus infections such as Epstein-Barr virus²⁶, simian virus 40^{27,28} and influenza virus^{29–31}. In this context it has been suggested that processed products rather than the intact protein may represent the actual antigen³¹ or that the intracellular protein may modify other viral or cellular products²⁹.

Irrespective of the so far unknown molecular identity of LYDIEA, the requirement of IE proteins for the synthesis of viral structural proteins during acute infection or after reactivation from latency raises the possibility of a physiological role of LYDIEA-specific CTL in the surveillance of viral latency. LYDIEA-specific memory-CTL which are present in latently infected mice⁸ could be activated as soon as IE genes are expressed and eliminate recurrently infected cells before the assembly of progeny virions.

This work was supported by two grants to U.H.K. of the Deutsche Forschungsgemeinschaft, SFB 120 and Ko 571/8. We thank Ms A. Lüske for technical assistance.

Received 17 April; accepted 11 September 1984.

1. Meyers, J. D., Flournoy, N. & Thomas, E. D. *Rev. infect. Dis.* **4**, 1119–1132 (1982).
2. Jordan, M. C., Shanley, J. D. & Stevens, J. G. *J. gen. Virol.* **37**, 419–423 (1977).
3. Mayo, D. R., Armstrong, J. A. & Ho, M. *Nature* **267**, 721–723 (1977).
4. Quinnan, G. V., Manischewitz, J. E. & Ennis, F. A. *Nature* **273**, 541–543 (1978).
5. Ho, M. *Infect. Immun.* **27**, 767–776 (1980).
6. Grundy, J. E. & Melief, C. J. M. *J. gen. Virol.* **61**, 133–136 (1982).
7. Borysiewicz, L. K., Morris, S., Page, J. D. & Sissons, J. G. P. *Eur. J. Immun.* **13**, 804–809 (1983).
8. Reddehase, M. J., Keil, G. M. & Koszinowski, U. H. *Eur. J. Immun.* **14**, 56–61 (1984).
9. Roizman, B. in *Perspectives in Virology* (ed. Pollard, M.) 283–301 (Harper and Row, New York, 1965).
10. Spear, P. G. & Roizman, B. *Cold Spring Harb. Monogr.* **10B**, 658–667 (1980).
11. Honess, R. W. & Roizman, B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1276–1280 (1975).
12. Imperiale, M. J., Feldman, L. T. & Nevins, J. R. *Cell* **35**, 127–136 (1983).
13. Green, M. R., Treisman, R. & Maniatis, T. *Cell* **35**, 137–148 (1983).
14. Watson, R. J. & Clements, J. B. *Nature* **285**, 329–330 (1980).
15. Stinski, M. F., Thomsen, D. R., Stenberg, R. M. & Goldstein, L. C. *J. Virol.* **46**, 1–14 (1983).
16. Stenberg, R. M., Thomsen, D. R. & Stinski, M. F. *J. Virol.* **49**, 190–199 (1984).
17. Marks, J. R., Mercer, J. A. & Spector, D. H. *Virology* **131**, 247–254 (1983).
18. Ebeling, A., Keil, G. M., Knust, E. & Koszinowski, U. H. *J. Virol.* **47**, 421–433 (1983).
19. Keil, G. M., Ebeling-Keil, A. & Koszinowski, U. H. *J. Virol.* **50**, 784–795 (1984).
20. Reddehase, M. J., Keil, G. M. & Koszinowski, U. H. *J. Immun.* **132**, 482–489 (1984).
21. Koszinowski, U. H., Gething, M.-J. & Waterfield, M. *Nature* **267**, 160–163 (1977).
22. Bevan, M. J. & Cohn, M. *J. Immun.* **114**, 559–565 (1975).
23. Berke, G. *Immun. Rev.* **72**, 5–42 (1983).
24. Chantler, J. K. & Hudson, J. B. *Virology* **86**, 22–36 (1978).
25. Rouse, B. T., Larsen, H. S. & Wagner, H. *Infect. Immun.* **39**, 785–792 (1983).
26. Rickinson, A. B., Wallace, L. E. & Epstein, M. A. *Nature* **283**, 865–867 (1980).
27. Gooding, L. R. & O'Connell, K. A. *J. Immun.* **131**, 2580–2586 (1983).
28. Tevethia, S. S., Tevethia, M. J., Lewis, A. J., Reddy, U. B. & Weissman, S. M. *Virology* **128**, 319–330 (1983).
29. Bennink, J. R., Yewdell, J. W. & Gerhard, W. *Nature* **296**, 75–76 (1982).
30. Townsend, A. R. M. & Skehel, J. J. *Nature* **300**, 655–657 (1982).
31. Townsend, A. R. M. & Skehel, J. J. *J. exp. Med.* **160**, 552–563 (1984).
32. Taswell, C. *J. Immun.* **126**, 1614–1619 (1981).
33. Fazekas de St. Groth, S. *J. Immun. Meth.* **49**, R11–R13 (1982).
34. Hudson, J. B., Misra, V. & Mosmann, T. R. *Virology* **72**, 235–243 (1976).
35. Shanley, J. D. *J. gen. Virol.* **63**, 251–254 (1982).

Phorbol ester and diacylglycerol mimic growth factors in raising cytoplasmic pH

W. H. Moolenaar, L. G. J. Tertoolen & S. W. de Laat

Hubrecht Laboratory, International Embryological Institute, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

There is now good evidence that cytoplasmic pH (pH_i) may have an important role in the metabolic activation of quiescent cells^{1,2}. In particular, growth stimulation of mammalian fibroblasts leads to a rapid increase in pH_i (refs 3–6), due to activation of a Na^+/H^+ exchanger in the plasma membrane⁴, and this alkalization is necessary for the initiation of DNA synthesis⁷. However, the mechanism by which mitogens activate the Na^+/H^+ exchanger to raise pH_i is not known, although an increase in cytoplasmic free Ca^{2+} ($[Ca^{2+}]_i$) has been postulated as the primary trigger^{8–10}. We

now present data suggesting that the Na^+/H^+ exchanger is set in motion through protein kinase C, a phospholipid- and Ca^{2+} -dependent enzyme normally activated by diacylglycerol produced from inositol phospholipids in response to external stimuli^{11,12}. Using newly developed pH microelectrodes and fluorimetric techniques, we show that a tumour promoting phorbol ester and synthetic diacylglycerol, both potent activators of kinase C (refs 12–15), mimic the action of mitogens in rapidly elevating pH_i in different cell types. Furthermore, we demonstrate that, contrary to previous views, an early rise in $[\text{Ca}^{2+}]_i$ is not essential for the activation of Na^+/H^+ exchange and the resultant increase in pH_i . Finally, we suggest that an alkaline pH_i shift, mediated by Na^+/H^+ exchange, may be a common signal in the action of those hormones which elicit the breakdown of inositol phospholipids.

Human foreskin fibroblasts (HF cells), HeLa cells and mouse N1E-115 neuroblastoma cells, grown on rectangular glass coverslips, were loaded with the fluorescent pH indicator 2',7'-bis(carboxyethyl)-5,6-carboxyfluorescein (BCECF) by uptake of its membrane-permeable ester¹⁶, and the pH_i -dependent emission intensity from the cells was continuously monitored as described previously^{4,17}. In addition, pH_i was directly measured in large N1E-115 cells by means of a novel type of pH-sensitive intracellular microelectrode (see Fig. 1 legend).

The tumour promoter 12-*O*-tetradecanoyl-phorbol-13-acetate (TPA), which binds to and directly activates protein kinase C^{11,12}, induces a rise in pH_i of ~ 0.15 units within 10 min in all three cell types studied. Typical pH_i recordings are shown in Fig. 1a, b, while Table 1 summarizes the major results. The effect of TPA on pH_i is dose dependent, half-maximal effect being observed at $\sim 25 \text{ ng ml}^{-1}$, whereas $0.5 \mu\text{g ml}^{-1}$ of the biologically inactive TPA analogue 4- α -phorbol-12,13-didecanoate fails to elevate pH_i (not shown). The direct intracellular recordings in N1E-115 cells also show that TPA has no significant effect on membrane potential (Fig. 1b).

The TPA-induced pH_i shifts resemble those induced by polypeptide growth factors in fibroblasts^{4–6}, suggesting that the underlying mechanism is the same, namely activation of Na^+/H^+ exchange. Indeed, TPA stimulates Na^+ influx in quiescent fibroblasts¹⁸, while the TPA-induced pH_i response of all three cell types is completely inhibited by removal of external Na^+ (Fig. 1c) and by $25 \mu\text{M}$ of 5-*N*-ethylisopropylamiloride (not shown), a potent inhibitor of the mammalian Na^+/H^+ exchanger (ref. 19 and W. H. M. and E. Cragoe, unpublished data). Furthermore, the effects of platelet-derived growth factor (PDGF) and TPA on pH_i in HF cells are not additive (Fig. 1d). Yet, the PDGF/TPA-induced pH_i shift is apparently not limited

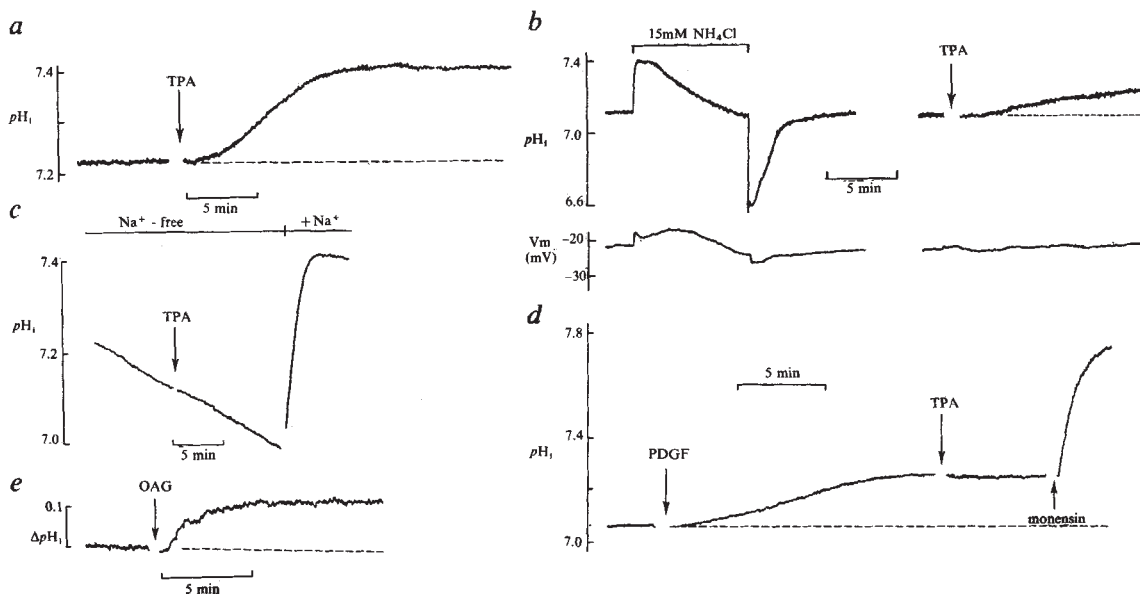


Fig. 1 *a*, Time course of increase in pH_i induced by TPA (200 ng ml^{-1}) in mouse N1E-115 neuroblastoma cells. Subconfluent cultures, attached to a glass coverslip, were deprived of serum for 24 h and loaded with BCECF as described previously⁴. The cells were incubated in HEPES-buffered Dulbecco's modified Eagle's medium (DMEM, pH 7.40) without phenol red and pH_i -dependent fluorescence was measured at 506–532 nm using a Perkin Elmer model 3000 fluorescence spectrometer. The fluorescence of BCECF varies linearly with pH over the range 6.4–7.6 (ref. 17). Temperature was maintained at 32°C . Similar TPA-induced shifts in pH_i were observed in serum-deprived HeLa cells (subconfluent) and HF cells (confluent). *b*, Direct measurement of pH_i and membrane potential (V_m) in large N1E-115 cells (diameter $> 40 \mu\text{m}$), penetrated by a pH-sensitive microelectrode and a voltage-measuring microelectrode, showing rapid recovery of pH_i from a NH_4^+ -induced acid load, and TPA-induced rise in pH_i . pH_i regulation in N1E-115 cells, as in HF cells and HeLa cells, is almost entirely mediated by amiloride-sensitive Na^+/H^+ exchange (refs 4, 17, 32 and W.H.M., unpublished results). The method for constructing pH-sensitive microelectrodes is described below. Solutions were exchanged by rapid perfusion. Other experimental details are essentially similar to those described previously^{33,34}. Final TPA concentration, 200 ng ml^{-1} . External pH 7.40. Temperature, 35°C . *c*, Lack of effect of TPA on pH_i in BCECF-loaded N1E-115 cells maintained in Na^+ -free medium. Na^+ was replaced by bis(2-hydroxyethyl)dimethylammonium. Cells rapidly acidify in Na^+ -free medium, but pH_i immediately recovers after re-addition of Na^+ (see refs 4, 17, 32). *d*, Effects of PDGF (40 ng ml^{-1}), TPA (200 ng ml^{-1}) and monensin ($5 \mu\text{M}$) on pH_i in BCECF-loaded culture of quiescent HF cells. Insulin was present at $5 \mu\text{g ml}^{-1}$ to potentiate PDGF-induced pH_i shift (insulin alone has no effect on pH_i ; ref. 4). Human PDGF ($>95\%$ pure) was added in the presence of $100 \mu\text{g ml}^{-1}$ bovine serum albumin (essentially fatty acid free). *e*, Effect of synthetic OAG ($50 \mu\text{g ml}^{-1}$) on pH_i in BCECF-loaded N1E-115 cells. OAG stock solution in dimethyl sulphoxide (DMSO) was diluted to 1% DMSO with DMEM + $50 \mu\text{g ml}^{-1}$ bovine serum albumin (fatty acid free). A $250\text{-}\mu\text{l}$ aliquot of this OAG solution was then briefly sonicated under N_2 and immediately added to the cells. The final DMSO concentration was 0.1%, which alone had no effect.

Methods: Construction of pH-sensitive microelectrodes: An 8-mm length of pH-sensitive glass tubing (outer diameter 1.0 mm; inner diameter 0.5 mm; Clark Electromedical Instruments) was placed in the centre of a 5-cm piece of insulating lead glass tubing (outer diameter 1.6 mm; wall thickness $< 0.15 \text{ mm}$). The coaxial tubes were pulled into two microelectrodes using a horizontal micropipette puller. About 50–60% of the electrodes then contains a sealed pH-sensitive tip with an outer diameter of $< 0.5 \mu\text{m}$. Electrodes were first filled with methanol by boiling under reduced pressure at room temperature. Methanol was then replaced with 0.1 M HCl by diffusion. Subsequently, the electrode tips were immersed into a chromic acid cleaning solution for 12–24 h before experimentation. Electrodes were tested for giving a nernstian response to pH changes ($54\text{--}58 \text{ mV per pH unit}$) and having resistance of $1\text{--}2 \times 10^9 \Omega$. The speed of response to a sudden pH change of 1 unit was about 15–20 s. Selected electrodes could be stored for a few days (tips in chromic acid) before obvious deterioration. The construction of this novel type of pH-sensitive electrode was found to be much easier than that of the common 'Thomas type' pH microelectrodes³⁵. Further details will be described elsewhere.

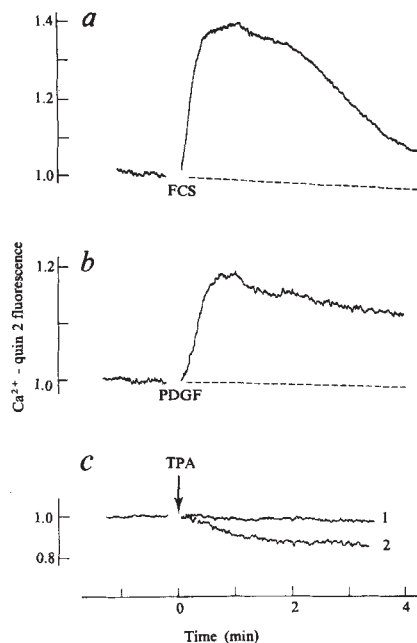


Fig. 2 Rapid changes in $[Ca^{2+}]_i$ in HF cells after addition of FCS (10%) (a), PDGF (40 ng ml⁻¹) (b) and TPA (c; trace 1, 50 ng ml⁻¹; trace 2, 200 ng ml⁻¹) at time zero. Quiescent HF monolayers were loaded with quin-2 by incubating them in DMEM containing 50 μ M quin-2 acetoxymethyl ester (Amersham) for 30 min at 37°C and $[Ca^{2+}]_i$ -dependent fluorescence was continuously measured at 339–500 nm (temperature, 32°C). Signals were calibrated essentially as described elsewhere^{26,36}. Resting value of $[Ca^{2+}]_i$ is estimated at 140 nM, while peak values of the FCS and PDGF traces correspond to ~400 and 300 nM $[Ca^{2+}]_i$, respectively.

by the magnitude of the transmembrane Na^+ and H^+ gradients, as pH_i can be increased still further by addition of the Na^+/H^+ ionophore monensin (Fig. 1d). From these results we conclude that TPA, like growth factors^{4–6}, raises pH_i by activating the otherwise quiescent Na^+/H^+ exchanger in the plasma membrane.

If activation of Na^+/H^+ exchange is mediated by kinase C, an alkaline pH_i shift may be expected after addition of an unsaturated diacylglycerol such as 1-oleoyl-2-acetyl-glycerol (OAG; refs 14, 15). In physiological conditions, 1,2-diacylglycerol is produced from hydrolysis of inositol phospholipids and serves as the endogenous activator of protein kinase C^{11,12}. As illustrated in Fig. 1e, addition of 50 μ g ml⁻¹ OAG to N1E-115 cells does indeed induce a rapid alkaline shift in pH_i of at least 0.1 unit. A similar shift was observed in OAG-treated HeLa cells. Other lipid degradation products such as phosphatidic acid and arachidonic acid, a precursor of prostaglandins, were unable to raise pH_i . Furthermore, indomethacin (5 μ M), a potent inhibitor of prostaglandin synthesis, did not affect the pH_i shifts induced by either TPA or PDGF. Various lipid-interacting drugs, such as the phenothiazines, are potent inhibitors of protein kinase C^{20,21}. We found that trifluoperazine (TFP) at 10 μ M produced ~50% inhibition of both the PDGF- and TPA-evoked pH_i shifts in HF cells without affecting basal pH_i levels. (At higher TFP doses the cytoplasm gradually acidified, presumably due to a nonspecific membrane perturbation and consequent H^+ influx.) The above results, together with the fact that growth factors are potent inducers of inositol lipid breakdown^{22,23}, strongly suggest that protein kinase C is responsible for the activation of Na^+/H^+ exchange by mitogens and phorbol ester.

Although the Ca^{2+} ionophore A23187 is incapable of elevating pH_i in fibroblasts⁴, the possibility nevertheless remains that a rapid increase in $[Ca^{2+}]_i$ may have some critical role in Na^+/H^+ exchange activation, as postulated by others¹⁰. We therefore used the intracellularly-trapped fluorescent dye quin-2 (refs 24,

Table 1 Effects of TPA on pH_i

Cell type	Resting pH_i (mean $\pm$ s.e.)	ΔpH_i (relative to control)
HF	7.05 $\pm$ 0.02 (n = 12)	0.16 $\pm$ 0.03 (n = 5)
HeLa	7.18 $\pm$ 0.04	0.17 $\pm$ 0.02 (n = 7)
N1E-115	7.24 $\pm$ 0.03	0.15 $\pm$ 0.02 (n = 8)

Cells were loaded with BCECF and assayed for TPA-induced shifts in pH_i (ΔpH_i) by the K^+ -nigericin calibration method as detailed elsewhere^{4,17}. HF and N1E-115 cells were maintained in serum-free DMEM for 24 h and HeLa cells for 4 h before experimentation. TPA concentration, 200 ng ml⁻¹. External pH , 7.30–7.40. Temperature, 30–32°C. The number of experiments is given in parentheses.

25) to monitor $[Ca^{2+}]_i$ after activation of quiescent HF and N1E-115 cells. Figure 2a, b illustrates typical results from HF cells, the details of which are presented elsewhere²⁶, and shows that fetal calf serum (FCS) and PDGF immediately raise $[Ca^{2+}]_i$ two- to threefold above its resting level of about 140 nM. This mitogen-induced Ca^{2+} signal is initiated without a detectable lag period and peaks within 1 min before slowly declining towards a new steady level. The Ca^{2+} response is not prevented by removal of external Ca^{2+} (ref. 26), indicating that FCS and PDGF act by releasing Ca^{2+} from intracellular stores, probably due to the rapid formation of inositol 1,4,5-trisphosphate²⁷. A qualitatively similar Ca^{2+} signal was observed in serum-stimulated N1E-115 cells (not shown).

By contrast, TPA, at concentrations that evoke a significant pH_i shift, completely fails to raise $[Ca^{2+}]_i$ above basal levels in HF cells (Fig. 2c). At 200 ng ml⁻¹, TPA even induces a slight fall in $[Ca^{2+}]_i$. The apparent inability of TPA to elevate $[Ca^{2+}]_i$ agrees with previous findings in quin-2-loaded lymphocytes²⁴, neutrophils²⁸ and platelets²⁹. These findings demonstrate that although growth factors are Ca^{2+} -mobilizing hormones, an early rise in $[Ca^{2+}]_i$ is not essential for the activation of Na^+/H^+ exchange by external stimuli, and they reinforce the notion that the Na^+/H^+ exchanger is set in motion mainly through activated kinase C. Thus, the simplest and most attractive explanation for our findings is that protein kinase C directly phosphorylates the Na^+/H^+ exchanger on its cytoplasmic face and thereby increases its apparent affinity for intracellular H^+ (see ref. 4).

The results reported here raise the interesting possibility that cytoplasmic alkalization, mediated by Na^+/H^+ exchange, is not restricted to the action of growth factors and tumour-promoting phorbol esters, but that it may commonly occur in response to those hormones and other surface stimulants which trigger the breakdown of inositol phospholipids and thereby activate kinase C. There is indeed some evidence to support this notion: in typical model systems of hormone-induced inositol lipid turnover, such as platelets and neutrophils, the pH_i rises rapidly on activation by physiological stimuli^{30,31}. Further work will be needed to explore the ubiquity of cytoplasmic alkalization in hormone action and to identify the pH_i -dependent responses in different target cells.

We thank Roger Y. Tsien for gifts of BCECF-ester and OAG, Carl-Henrik Heldin for providing pure PDGF, Lilian Joosen for technical assistance and Dr M. J. Berridge for stimulating discussions. This work was supported by grants from the Netherlands Cancer Foundation (Koningin Wilhelmina Fonds).

Received 15 May; accepted 11 September 1984.

- Nuccitelli, R. & Deamer, D. W. (eds) *Intracellular pH: Its Measurement, Regulation and Utilization in Cellular Functions* (Liss, New York, 1982).
- Busa, W. B. & Nuccitelli, R. *Am. J. Physiol.* **246**, R409–R438 (1984).
- Schuldiner, S. & Rozengurt, E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7778–7782 (1982).
- Moolenaar, W. H., Tsien, R. Y., van der Saag, P. T. & de Laat, S. W. *Nature* **304**, 645–648 (1983).
- Cassel, D., Rothenberg, P., Zhuang, Y., Deuel, T. F. & Glaser, L. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6224–6228 (1983).
- L'Allemain, G., Paris, S. & Pouyssegur, J. *J. biol. Chem.* **259**, 5809–5815 (1984).
- Pouyssegur, J., Sardet, C., Franchi, A., L'Allemain, G. & Paris, S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4833–4839 (1984).
- Villereal, M. L. *J. cell. Physiol.* **107**, 359–369 (1981).
- Benos, D. J. *Am. J. Physiol.* **242**, C131–C145 (1982).

10. Owen, N. E. & Villereal, M. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3537-3541 (1982).
11. Takai, Y., Kishimoto, A. & Nishizuka, Y. in *Calcium and Cell Function* Vol. 2 (ed. Cheung, W. Y.) 385-412 (Academic, New York, 1982).
12. Nishizuka, Y. *Nature* **308**, 693-698 (1984).
13. Castagna, M. *et al. J. biol. Chem.* **257**, 7847-7851 (1982).
14. Mori, T. *et al. J. Biochem.* **91**, 427-431 (1982).
15. Kaibuchi, K. *et al. J. biol. Chem.* **258**, 6701-6704 (1983).
16. Rink, T. J., Tsien, R. Y. & Pozzan, T. *J. Cell Biol.* **95**, 189-196 (1982).
17. Moolenaar, W. H., Tertoolen, L. G. J. & de Laat, S. W. *J. biol. Chem.* **259**, 7563-7570 (1984).
18. Dicker, P. & Rozengurt, E. *Biochem. biophys. Res. Commun.* **100**, 433-441 (1981).
19. Vigne, P., Frelin, C., Cragoe, E. & Lazdunski, M. *Molec. Pharmacol.* **25**, 131-136 (1984).
20. Mori, T., Takai, Y., Minakuchi, R., Yu, B. & Nishizuka, Y. *J. biol. Chem.* **255**, 8378-8380 (1980).
21. Wise, B. C. *et al. J. biol. Chem.* **257**, 8489-8495 (1982).
22. Habenicht, A. *et al. J. biol. Chem.* **256**, 12329-12335 (1981).
23. Sawyer, S. T. & Cohen, S. *Biochemistry* **20**, 6280-6286 (1981).
24. Tsien, R. Y., Pozzan, T. & Rink, T. J. *Nature* **295**, 68-71 (1982).
25. Tsien, R. Y., Pozzan, T. & Rink, T. J. *J. Cell Biol.* **94**, 325-334 (1982).
26. Moolenaar, W. H., Tertoolen, L. G. J. & de Laat, S. W. *J. biol. Chem.* **259**, 8066-8069 (1984).
27. Berridge, M. J., Heslop, J. P., Irvine, R. F. & Brown, K. D. *Biochem. J.* **222**, 195-201 (1984).
28. Sha'afi, R. I. *et al. Biochem. biophys. Res. Commun.* **114**, 638-645 (1983).
29. Rink, T. J., Sanchez, A. & Hallam, T. J. *Nature* **305**, 317-319 (1983).
30. Molski, T. *et al. Biochem. biophys. Res. Commun.* **94**, 508-514 (1980).
31. Horne, W. C., Norman, N. E., Schwartz, D. B. & Simons, E. R. *Eur. J. Biochem.* **120**, 295-302 (1981).
32. Moolenaar, W. H., Boonstra, J., van der Saag, P. T. & de Laat, S. W. *J. biol. Chem.* **256**, 12883-12887 (1981).
33. Moolenaar, W. H. & Spector, I. *J. Physiol., Lond.* **278**, 265-286 (1978).
34. Moolenaar, W. H., Mummery, C. L., van der Saag, P. T. & de Laat, S. W. *Cell* **23**, 789-798 (1981).
35. Thomas, R. C. *J. Physiol., Lond.* **238**, 159-180 (1974).
36. Hesketh, T. R., Smith, G. A., Moore, J. P., Taylor, M. V. & Metcalfe, J. C. *J. biol. Chem.* **258**, 4876-4882 (1983).

Stepwise enzymatic dephosphorylation of inositol 1,4,5-trisphosphate to inositol in liver

Diane J. Storey, Stephen B. Shears, Christopher J. Kirk & Robert H. Michell

Department of Biochemistry, University of Birmingham, PO Box 363, Birmingham B15 2TT, UK

Many receptors for hormones, neurotransmitters and other signals cause hydrolysis of phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P₂) and effect a rise in cytosolic Ca²⁺ concentration¹⁻⁷. The inositol 1,4,5-trisphosphate (Ins(1,4,5)P₃) liberated during PtdIns(4,5)P₂ breakdown seems to serve as a second messenger that activates the release of Ca²⁺ from a nonmitochondrial intracellular compartment⁷⁻¹². As expected if it is an important intracellular messenger, Ins(1,4,5)P₃ is relatively rapidly degraded, both within stimulated cells^{13,14} and when added to homogenates of blowfly salivary gland¹⁵ or to permeabilized, but not intact, hepatocytes¹⁰. Here we report that the dephosphorylation reactions responsible for the conversion of Ins(1,4,5)P₃ to free inositol in rat liver are catalysed by two or more enzymes, and that these reactions are distributed between the plasma membrane and cytosol. The Ins(1,4,5)P₃ 5-phosphatase and inositol 1-phosphate (Ins(1)P) phosphatase of liver appear similar to enzymes described previously in erythrocytes¹⁶ and brain¹⁷.

Rat liver homogenates were incubated with small concentrations of ³²P-labelled inositol phosphates in conditions of ionic composition and pH chosen to approximately mimic those of the cytosol of stimulated cells (for details, see legend to Table 1). Ins(1,4,5)P₃ (labelled in its 4- and 5-phosphate groups), inositol-1,4-bisphosphate(Ins(1,4)P₂) (labelled in its 4-phosphate) and Ins(1)P were rapidly degraded during such incubations, with hydrolysis of Ins(1,4,5)P₃ being the most rapid reaction (Table 1). High-speed centrifugation sedimented a large proportion of the Ins(1,4,5)P₃ phosphatase activity and left most of the Ins(1,4)P₂ phosphatase and Ins(1)P phosphatase activity in the soluble fraction (Table 1). Analysis of the products of Ins(1,4,5)P₃ hydrolysis^{16,18} by a plasma membrane/Golgi-rich membrane fraction recovered from a density gradient separation (as in Fig. 1) showed selective removal of the 5-phosphate (data not shown).

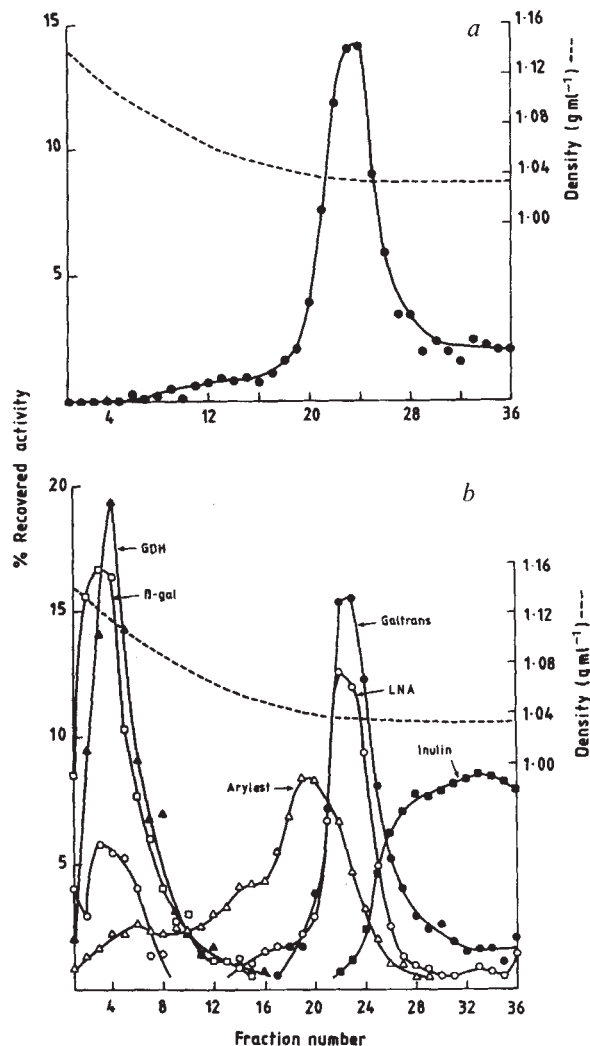


Fig. 1 Distribution of Ins(1,4,5)P₃ 5-phosphatase and various marker enzymes after density gradient centrifugation of a rat liver nuclei-free homogenate. *a*, The distribution of Ins(1,4,5)P₃ 5-phosphatase. *b*, The distribution of marker enzymes for cytosol (inositol), Golgi membranes (galactosyltransferase, galtrans), plasma membranes (leucyl- β -naphthylamidase, LNA), endoplasmic reticulum (arylesterase, arylest), lysosomes (β -galactosidase, β -gal) and mitochondria (glutamate dehydrogenase, GDH). The results shown and the recoveries (see below) are for a single representative experiment. Data are plotted as the percentages of the total recovered activity or component found in each fraction, and values of <0.5% of recovered activity have been omitted from *b* for clarity. **Methods:** Hepatocytes, prepared and incubated as described previously²⁴, were homogenized using N₂ cavitation²⁵ in 0.25 M sucrose containing 3 mM TES (2-[(2-hydroxy-1,1-bis(hydroxymethyl)ethyl)amino]ethane sulphonic acid), adjusted to pH 7.4 with KOH; 0.4 ml of a nuclei-free homogenate²⁵ was added to 0.4 ml of homogenization buffer that had been layered on the top of an isoosmotic density gradient of Percoll suspended in homogenization buffer (5.5 ml of density 1.052 g ml⁻¹ above 1.2 ml of density 1.074 g ml⁻¹). Gradients were centrifuged for 15 min at 73,000g (mean) in a 10 $\times$ 10 ml angle rotor. Fractions (214 μ l) were recovered, and their components and densities assayed as follows: Ins(1,4,5)P₃ 5-phosphatase (assayed as described in Table 1; recovery of loaded activity in fractions from gradient-85%); inulin (followed in a separate gradient in parallel, after addition of 1.5 μ Ci ml⁻¹ ³H-inulin to homogenate before centrifugation; 100% recovery); galactosyltransferase (ref. 24; 84% recovery); LNA (ref. 25; 96% recovery) (a large proportion of this enzyme activity distributes with glucagon-stimulated adenylate cyclase, an enzyme concentrated in the hormone-sensitive sinusoidal domain of the hepatocyte plasma membrane²⁵); arylesterase (ref. 24; 96% recovery); β -galactosidase (ref. 26; 111% recovery); GDH (ref. 24; 81% recovery).

Table 1 First-order rate constants for hydrolysis of inositol phosphates in liver preparations, corrected to the enzyme concentrations of the original tissue

Fraction	Ins(1,4,5)P ₃	First-order rate constant (s ⁻¹)			Ins(4)P
		Ins(1,4)P ₂	Ins(1)P		
Homogenate	0.148 ± 0.019	0.018 ± 0.022	0.007 ± 0.001		NA
Soluble	0.035 ± 0.005	0.028 ± 0.004	0.009 ± 0.002		0.0002
Particulate	0.095 ± 0.007	0.004 ± 0.001	0.003 ± 0.0004		NA
Washed particulate	0.065 ± 0.006	0.004 ± 0.0004	0.0009 ± 0.0003		NA

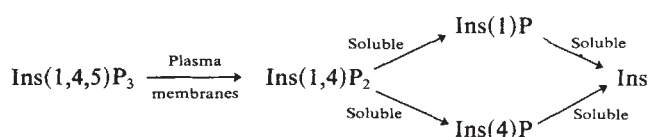
[4-³²P]Ins(1,4)P₂ and [4,5-³²P]Ins(1,4,5)P₃ were prepared as described previously^{15,17}. ³²P-Ins(1)P was produced by phosphodiesterase-catalysed hydrolysis of ³²P-phosphatidylinositol using a soluble liver enzyme preparation (details to be described elsewhere), and isolated by high-voltage electrophoresis²³. Rat livers were cleared of blood by perfusion of 0.25 M sucrose/5 mM Tris (pH 8.0) through the hepatic portal vein under pentobarbital anaesthesia. The livers were removed, homogenized in the same medium in a glass-Teflon Potter-Elvehjem homogenizer (20 strokes, 500 r.p.m.), and the homogenates filtered through a 300-μm mesh and made up to 3 ml per g tissue. Homogenates were centrifuged at 180,000g for 30 min and the supernatant and pellet made up to the volume of the original homogenate (to yield 'soluble' and 'particulate' fractions). Some of the 'particulate' fraction was again centrifuged and resuspended, to yield the 'washed particulate' fraction. The medium used for enzyme incubations comprised (in mM): 10 NaCl, 70 K⁺ glutamate, 30 KCl, 4 MgSO₄, 10 HEPES, 1 EGTA, pH 7.2. Ins(1,4,5)P₃ 5-phosphatase activity in this 'physiological' medium is rather more than half of that in the ionic strength medium used previously¹⁶, both for erythrocyte membranes and liver cell fractions. Incubations were in a final volume of 1 ml, containing 50 μl of the homogenate or cell fraction, and they were continued for an appropriate period (1 min to 4 h) in order to yield a first-order rate constant for the hydrolysis of the phosphate ester being tested. The substrate concentrations were in the ranges 40–300 nM (for Ins(1,4,5)P₃ and Ins(1,4)P₂) and 8–24 μM (for Ins(1)P). Incubations were stopped by adding 100 μl of 12 M perchloric acid containing 1 mM P_i, and released ³²P-P_i was estimated as described previously¹⁶. The values given are the means ± s.e.m. (n = 4). The approximate rate constant for hydrolysis of Ins(4)P is derived from the second exponential (that for ³²P-Ins(4)P) on a prolonged rate curve describing hydrolysis of [4-³²P]Ins(1,4)P₂. NA, not assayed.

On density gradient centrifugation of nuclei-free liver homogenates, the Ins(1,4,5)P₃ 5-phosphatase activity of the particulate fraction was largely associated with fractions of density 1.035 g ml⁻¹; these were rich in plasma membrane and Golgi markers, rather than those of mitochondria, endoplasmic reticulum or lysosomes (Fig. 1). Isolated nuclei showed little Ins(1,4,5)P₃ phosphatase activity (data not shown), and Prentki *et al.*¹¹ have also reported a low Ins(1,4,5)P₃ phosphatase activity in endoplasmic reticulum membranes from an insulinoma-derived cell line. Preliminary studies of highly enriched preparations of plasma membranes and Golgi membranes suggest that most of the Ins(1,4,5)P₃ 5-phosphatase is associated with plasma membranes, as in the erythrocyte¹⁶. The membrane-bound activity in liver was near-maximal over a fairly broad range of pH (~5.0–7.2), also like the erythrocyte enzyme. The erythrocyte Ins(1,4,5)P₃ 5-phosphatase is very sensitive to inhibition by Ag⁺ and Zn²⁺ (ref. 16), but problems of solubility prevented the use of these ions in the iso-ionic media employed in our experiments with liver cell fractions. However, we have recently shown that the erythrocyte enzyme is very sensitive to inhibition by Cd²⁺ (C. E. King, unpublished data), and the membrane-bound Ins(1,4,5)P₃ phosphatase of liver is similarly inhibited by this ion (Fig. 2). The membrane-bound Ins(1,4,5)P₃ 5-phosphatase activity was insensitive to concentrations of Li⁺ up to 10 mM and was inhibited only by high concentrations of F⁻ (k_i ~3 mM). The smaller proportion of the Ins(1,4,5)P₃ phosphatase activity that was present in the soluble fraction was studied in less detail (because of the problems inherent in assaying Ins(1,4,5)P₃ hydrolysis without interference from further hydrolysis of the liberated Ins(1,4)P₂), but its general characteristics appeared similar to those of the particulate activity.

The soluble activity that hydrolysed Ins(1,4)P₂ yielded both a labelled inositol monophosphate and labelled P_i. Because the Ins(1,4)P₂ used as substrate was labelled only in the 4-phosphate, initial attack must have occurred at either phosphate group. The labelled Ins(4)P produced during these incubations was slowly dephosphorylated. Ins(1,4)P₂ hydrolysis was much less sensitive to Cd²⁺ inhibition than was hydrolysis of Ins(1,4,5)P₃ (Fig. 2); it was, however, sensitive to inhibition by Li⁺ (k_i ~2.5 mM) and F⁻ (k_i ~0.2 mM). The soluble activity responsible for Ins(1)P hydrolysis, like the Ins(1)P phosphatase of brain¹⁷ and other tissues¹⁹, was sensitive to inhibition by Li⁺ (k_i ~4 mM) and F⁻ (k_i ~0.2 mM).

These results, obtained using incubation conditions similar to those within the cytosol, indicate that the pathways that inactivate Ins(1,4,5)P₃ and return its inositol to the free inositol

pool of the cell are as shown below:



These results have several interesting implications. First, hydrolysis of Ins(1,4,5)P₃ involves the selective removal of its 5-phosphate by a phosphatase which appears to be localized, at least in part, at the plasma membrane. This enzyme will effectively terminate the mobilization of Ca²⁺ by Ins(1,4,5)P₃, as it yields Ins(1,4)P₂ which is essentially incapable of mobilizing Ca²⁺ from intracellular stores^{7–12}. From the relatively limited results obtained so far, it seems that the Ins(1,4,5)P₃ 5-phosphatase of liver is very similar to that of the erythrocyte both in its plasma membrane localization and its inhibition characteristics. The function of the Ins(1,4,5)P₃ 5-phosphatase activity of erythrocyte plasma membranes is unknown, but it might be a remnant of a hormonal signalling system which functioned during the haematopoietic stages of the cells' development¹⁶. This would be analogous to the retention of the guanyl nucleotide-dependent coupling proteins of the adenylate cyclase system in mature erythrocytes (for example ref. 20). The similarity between the membrane-bound Ins(1,4,5)P₃ 5-phosphatases of liver and erythrocytes provides some support for the notion that the experimentally accessible erythrocyte enzyme may be useful for investigating the characteristics of the enzymatic mechanism(s) by which Ins(1,4,5)P₃ is inactivated in stimulated cells.

Hydrolysis of Ins(1,4)P₂ and of both isomers of InsP is much slower than hydrolysis of Ins(1,4,5)P₃, as also observed with blowfly salivary gland¹⁵. Thus it seems likely that under normal circumstances the former reactions control the rate at which inositol derived from Ins(1,4,5)P₃ is returned to the cellular pool of free inositol. In the presence of Li⁺, these reactions are substantially inhibited, and inositol is trapped in the phosphates^{19,21}.

The hydrolysis of Ins(1,4)P₂ by soluble phosphatase activity does not show a precise positional selectivity like that of Ins(1,4,5)P₃: the 1- and 4-phosphate groups are removed at approximately similar rates, and the released Ins(1)P and Ins(4)P are both more slowly dephosphorylated by soluble phosphatase(s). The studies reported here do not show whether a single enzyme, which is sensitive to inhibition by Li⁺ and F⁻, is responsible for hydrolysis of Ins(1,4)P₂ and of the inositol monophosphates, or whether a different enzyme is involved in

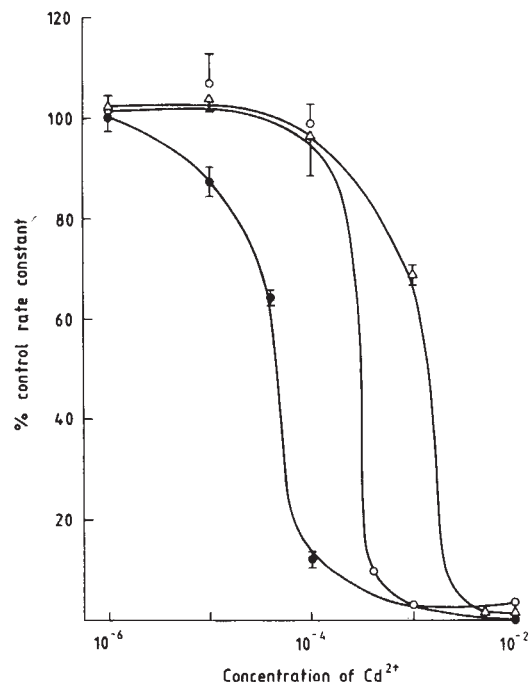


Fig. 2 A comparison of the sensitivities to inhibition by Cd²⁺ of particle-bound Ins(1,4,5)P₃ 5-phosphatase (●), soluble Ins(1,4)P₂ phosphatase (○) and Ins(1)P phosphatase (△) from rat liver. Enzyme activities were measured as described in Table 1, in the presence of various concentrations of Cd²⁺. Results are the mean $\pm$ s.e.m. ($n \geq 4$). Values for s.e.m. < 1% have been omitted.

each reaction. Sherman and colleagues²¹ have recently identified small quantities of Ins(4)P among the inositol phosphates that accumulate as a result of receptor activity in the brains of Li⁺-treated rats. It seems probable that this novel phosphate ester is produced from Ins(1,4,5)P₃, generated as a result of receptor activity, by the reactions shown in the scheme above.

An approximate extrapolation from the first-order rate constant given in Table 1 suggests that, in intact liver cells, the half life of Ins(1,4,5)P₃ is likely to be of the order of 4 s, suggesting considerably faster degradation of this compound than was reported in parotid fragments on termination of a stimulus^{13,14}. However, a recent report from Irvine and colleagues has shown that not only Ins(1,4,5)P₃, but also inositol 1,3,4-trisphosphate (Ins(1,3,4)P₃), accumulates in stimulated rat parotid fragments²², with the unexpected 1,3,4-isomer predominating. The reason only small quantities of Ins(1,4,5)P₃ were detected might be that this phosphate ester is degraded much more quickly than Ins(1,3,4)P₃; nothing is yet known about the pathway by which the latter compound is degraded.

We thank the MRC for project grant support and for a research studentship (to D.J.S.), and Eleanor Metcalf for her participation in the initial studies. We thank Alan Wallace for helpful advice on density gradient centrifugation.

14. Aub, D. & Putney, J. W. *Life Sci.* **34**, 1347-1355 (1983).
15. Berridge, M. J., Dawson, R. M. C., Downes, C. P., Heslop, J. P. & Irvine, R. F. *Biochem. J.* **212**, 473-482 (1983).
16. Downes, C. P., Mussat, M. C. & Michell, R. H. *Biochem. J.* **203**, 169-177 (1982).
17. Hallcher, L. M. & Sherman, W. R. *J. Biol. Chem.* **255**, 10896-10901 (1980).
18. Hawkins, P. T., Michell, R. H. & Kirk, C. J. *Biochem. J.* **218**, 785-793 (1984).
19. Berridge, M. J., Downes, C. P. & Hanley, M. R. *Biochem. J.* **206**, 587-595 (1982).
20. Kaslow, H. R., Johnson, G. L., Brothers, V. M. & Bourne, H. R. *J. Biol. Chem.* **255**, 3736-3741 (1980).
21. Sherman, W. R., Munsell, L. Y. & Honchar, M. P. *Chilton Conf.*, Dallas, 9-11 January, 1984 (in the press).
22. Irvine, R. F., Letcher, A. J., Lander, D. J. & Downes, C. P. *Biochem. J.* **223**, 237-243 (1984).
23. Dawson, R. M. C. & Clarke, R. *Biochem. J.* **127**, 113-118 (1972).
24. Shears, S. B. & Kirk, C. J. *Biochem. J.* **219**, 375-382 (1984).
25. Heyworth, C. M., Wallace, A. V. & Houslay, M. D. *Biochem. J.* **214**, 99-110 (1983).
26. Heyworth, C. M., Neumann, G. F. & Wynn, C. H. *Biochem. J.* **193**, 773-779 (1981).

A developmental gene product of *Bacillus subtilis* homologous to the sigma factor of *Escherichia coli*

Patrick Stragier*, Jean Bouvier*, Céline Bonamy† & Jekisiel Szulmajster†

* Institut de Microbiologie, Bâtiment 409, Université Paris-Sud, 91405 Orsay Cedex, France

† Laboratoire d'Enzymologie du CNRS, 91190 Gif-sur-Yvette, France

Sporulation of *Bacillus subtilis* involves sequential morphological and biochemical changes and is regulated by specific genes (*spo* genes) estimated to occupy more than 30 loci^{1,2}. A mutation in any one of these genes blocks the sporulation process at the corresponding developmental stage. Despite intensive genetic studies, the nature and function of the *spo* gene products remain unknown. Vegetative *B. subtilis* RNA polymerase core enzyme may interact with several sigma factors and discriminate among different classes of promoters³. During sporulation, new polypeptides are associated with the core enzyme which may have a central role in modifying its promoter recognition specificity^{3,4}. As a first step to understanding their function in the switch from vegetative to sporulation mode, several early sporulation genes have been cloned and analysed⁵⁻⁷. Here we report the cloning and nucleotide sequence of the *spoIIIG* gene of *B. subtilis*^{8,9}. This gene encodes a polypeptide with a predicted relative molecular mass of 27,652 which contains a 65-amino acid region highly homologous to an internal part of the *Escherichia coli* sigma factor.

From a genomic library constructed as described previously¹⁰, but enriched in DNA restriction fragments of 2-6 kilobases (kb), we isolated a new recombinant plasmid that was able to transform to the Spo⁺ phenotype both Rec⁺ and Rec⁻ *spoIIIG* recipients, indicating that the insert carried the entire *spoIIIG* locus. The 3.6-kb insert was analysed by endonuclease restriction mapping (Fig. 1) and was further reduced by several subcloning steps. Ligation of a 1.1-kb *Pst*I fragment in the *Pst*I site of the pHV33 vector yielded plasmid pGSIIG11, still carrying the intact *spoIIIG* locus, as shown by its transforming activity.

The 1,133-base pair (bp) nucleotide sequence of the pGSIIG11 insert (Fig. 2) was determined by the method of Maxam and Gilbert¹¹ according to the strategy outlined in Fig. 1. There is only one large open reading frame potentially encoding a polypeptide. As the requirements for translation initiation are more stringent in Gram-positive bacteria than in *E. coli*¹², a strong ribosome binding site is usually found in the vicinity of the translational start codon of *B. subtilis* genes. Such a sequence, AGAAGGGAG, appears at position 164 and precedes by six nucleotides an in-frame ATG codon. It seems probable that initiation of translation *in vivo* occurs at this codon, leading to the synthesis of a 239-amino acid polypeptide with a relative molecular mass (M_r) of 27,652. This polypeptide is rich in charged amino acids (33.5%) with an excess of basic residues over acidic ones (42 versus 38). These charged residues are not randomly distributed: the basic ones are clustered in the terminal

Received 27 July; accepted 7 September 1984.

1. Michell, R. H. *Biochim. biophys. Acta* **415**, 81-147 (1975).
2. Michell, R. H. *Trends biochem. Sci.* **4**, 128-131 (1979).
3. Michell, R. H., Kirk, C. J., Jones, L. M., Downes, C. P. & Creba, J. A. *Phil. Trans. R. Soc. B* **296**, 123-133 (1981).
4. Downes, C. P. & Michell, R. H. *Cell Calcium* **3**, 429-502 (1982).
5. Michell, R. H., Hawkins, P. T., Palmer, S. & Kirk, C. J. in *Calcium Regulation in Biological Systems* (ed. Endo, M.) (Takeda Foundation, Kyoto, in the press).
6. Berridge, M. J. *Molec. cell. Endocr.* **24**, 115-140 (1981).
7. Berridge, M. J. *Biochem. J.* **220**, 345-360 (1984).
8. Streb, H.-D., Irvine, R. F., Berridge, M. J. & Shulz, I. *Nature* **306**, 67-79 (1983).
9. Burgess, G. M. *et al. Nature* **309**, 63-66 (1984).
10. Joseph, S. K., Thomas, A. P., Williams, R. J., Irvine, R. F. & Williamson, J. R. *J. Biol. Chem.* **259**, 3077-3080 (1984).
11. Prentki, M. *et al. Nature* **309**, 562-564 (1984).
12. Irvine, R. F., Brown, K. D. & Berridge, M. J. *Biochem. J.* **222**, 269-272 (1984).
13. Downes, C. P. & Wusteman, M. *Biochem. J.* **216**, 633-640 (1983).

100
 CTCGAGATAAATGTTTAATCGGTATCAGCACAACAAGCTGTCGGCAGATGAGAGTTTGACGCGATTATTCATCCGAAATGCTTTCGGCGCAAGCGTGTCAA
 150
 ACACGTTTCATAATGTTCGCAATGTCCGTTACTTATCATACTCAAACGGCGTGACATTTAGAAGGAGAGGAAGATGAAAAAACTGAAATTACGGTTGACGCAC
 200
 fMetLysLysLysLeuArgLeuThrHis
 250
 CTCTGGTATAAGCTGCTGATGAACTTGGGCTGAAAAGTGATGAAGTCTATTACATAGCGGGAGTGAAGCCCTGCCGCCCTCCATTATCTAAAGATGAGGAGCAG
 300
 LeuTrpTyrLysLeuLeuMetLysLeuGlyLeuLysSerAspGluValTyrTyrIleGlyGlySerGluAlaLeuProProLeuSerLysAspGluGluGln
 350
 GTTTTGTTAATGAAGCTCCCAACGGCGATCAGCGCGCGCCATTCTAATTGAACGCAATTTGCGTCTGGTCGTATATATCGCCCGTAAATTTGAAATACG
 400
 ValLeuLeuMetLysLeuProAsnGlyAspGlnAlaAlaArgAlaIleLeuIleGluArgAsnLeuArgLeuValValTyrIleAlaArgLysPheGluAsnThr
 450
 GGAATTAATATAGAGGATTTAATCAGCATCGGTACCATCGGTCTAATCAAAGCTGTTAATACATTAAATCCAGAAAGAAAAATCAAGCTTGCTACCTATGCCTCC
 500
 GlyIleAsnIleGluAspLeuIleSerIleGlyThrIleGlyLeuIleLysAlaValAsnThrPheAsnProGluLysLysIleLysLeuAlaThrTyrAlaSer
 550
 CGGTGTATAGAAAATGAAATCTGATGTATTTAAGAAGAAATAACAAGAAATCCGTTCAAGAGTTTCCTTTGATGAACCGCTTAATATTGATTGGGACGGCAATGAG
 600
 ArgCysIleGluAsnGluIleLeuMetTyrLeuArgArgAsnAsnLysIleArgSerGluValSerPheAspGluProLeuAsnIleAspTrpAspGlyAsnGlu
 650
 CTTTGTCTTCTGATGTGCTCGGCACTGACGATGACATCATCACTAAGACATAGAAGCTAAGCTCGATAAAAAGCTTTTGAAAAAGCGCTTGAAACAGCTTAAT
 700
 LeuLeuLeuSerAspValLeuGlyThrAspAspAspIleIleThrLysAspIleGluAlaAsnValAspLysLysLeuLeuLysLysAlaLeuGluGlnLeuAsn
 750
 GAGAGAGAAAAGCAATCATGGAGCTCGCGTTTGGGCTTGTGCGTGAAGAAGAAAAACCCAAAGGATGTAGCGGATATGATGGGGATTCTCAGICTTATATT
 800
 GluArgGluLysGlnIleMetGluLeuArgPheGlyLeuValGlyGluGluGluLysThrGlnLysAspValAlaAspMetMetGlyIleSerGlnSerTyrIle
 850
 TCGCGGCTTGAGAAAAAGTAATAAAAAGGTGAGAAAAAGTTCACAAATGGTGTA AAAATTTTATGGTTAGAACCCCTTGATTTTACAGGATTTCCTGA
 900
 SerArgLeuGluLysArgIleIleLysArgLeuArgLysGluPheAsnLysMetVal●●●
 950
 TTTCGACAGTTTTTCGGTCTGAGTGCAGTGCATATTTTCCCACCCAGGAGATACCTAACGTTGTACAGCAGCTCCTGTAGGGAGGGAAAAAGTGTGAGAAA
 1000
 1100
 TAAAGTCGAAATCTCGGGGGTGATACCTCCAATTACCAGTACTCAAGAATGAAGAGATGAGAAAGCTGTTTAGGCAGCTGCAG

a

<i>B.subtilis</i> spoIIg	58	A	R	A	I	L	I	E	R	N	L	R	L	V	V	I	A	R	K	F	E	N	T	G	I	N	I	E	D	L	I	S	I	
<i>E.coli</i> sigma	375	A	K	K	E	M	V	E	A	N	L	R	L	V	I	S	I	A	K	K	Y	T	N	R	G	L	Q	F	L	D	L	I	Q	E

b

<i>B.subtilis</i> spoIIg	91	G	T	I	G	L	I	K	A	V	N	T	F	N	P	E	K	K	I	K	L	A	T	Y	A	S	R	C	I	E	N	E	I	122
<i>E.coli</i> sigma	408	G	N	I	G	L	M	K	A	V	D	K	F	E	Y	R	R	G	Y	K	F	S	T	Y	A	T	W	W	I	R	Q	A	I	439

b

<i>B.subtilis</i> spoIIg	350	GCGCGGCCCATCTCTAATTGAACGCAATTTGCGCTGGT	CGTATATATCGCCCGTAATTTGAAATACGGGAATTAATATAGAGGATTTAATCAGCATC																														
		***	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *
<i>E.coli</i> rpoD	1647	GCGAAGAAAGAGATGTTGAAGCGAACTTACGCTG	TGTTATTCTATCGCTAAGAAATACACCAACCGTGCTTGAGTTCCCTTGACCTGATTCAGGA																														

B.subtilis spoIIg 449 GGTACCATCGGTCTAATCAAGCTGTTAATACATTTAATCCAGAAAAGAAAATCAAGCTTGCTACCTATGCCTCCGGTGATAGAAAATGAAATC 544

E.coli rpoD 1746 GGCACACATCGGTCTGATGAAGCGGTTGATAAATTCGAATACCGCCGTGGTTACAAGTTCTCCACCTACGCAACCTGGTGGATCCGTGAGCGGCATC 1841

©1984Nature Publishing Group

parts of the polypeptide (8 among the first 22 residues and 8 among the last 18 residues), while a stretch of 10 acidic residues is found in an internal region, between amino acids 135 and 163. These features may be significant for the biological activity of the *spoIIG* gene product.

The encoded protein seems to be synthesized in a very low amount as we have been unable to express it in *B. subtilis* maxicells in various experimental conditions; this suggests that the protein may have a regulatory role in sporulation. Note that the *spoIIG* gene is transcribed only in the early stages of sporulation and not during vegetative growth. This was demonstrated by hybridization experiments in which the gene was used to probe mRNAs isolated at different stages of growth and sporulation (results not shown).

To elucidate the nature of the polypeptide encoded by the *spoIIG* gene, we searched for homologies with known proteins stored in the Dayhoff protein sequence data bank (~2,300 sequences). An important homology was found with the product of the *E. coli rpoD*¹³ gene, the sigma factor. Figure 3a shows the alignments of the two related segments of the *B. subtilis spoIIG* gene product and of the *E. coli* sigma factor. These aligned sequences exhibit a remarkable similarity: of the 65 positions compared, 29 (44.6%) are occupied by identical amino acids. When chemically similar pairs of amino acids are taken into account¹⁴, this number rises to 44 (67.7%). It was not necessary to introduce gaps to maximize this homology. Furthermore, statistical tests^{15,16} show that the similarities between the aligned sequences are highly significant; the probability that such a similarity is fortuitous is $<10^{-9}$ (Fig. 3a).

Analysis of the corresponding nucleotide sequences also revealed a high degree of similarity at the DNA level (100 matches over 195 positions, that is 51.3%; see Fig. 3b). A striking feature of the similarities is their clustering into three blocks of 34, 35 and 30 bp, where 73.5%, 80% and 70% of the respective positions are occupied by identical nucleotides. As the codon usage is not the same in *E. coli* and in *B. subtilis* (as indicated, for example, by the respective usages of the *trpE* genes of these two organisms^{17,18}), such similarities strongly suggest that the two genes are evolutionarily related.

The regions shown in Fig. 3 correspond to different internal parts of the two proteins: residues 58–127 of the 239-amino acid polypeptide encoded by *spoIIG* and residues 375–439 of the *E. coli* sigma factor, a polypeptide of 613 amino acids¹³. However, their high degree of similarity strongly suggests a common function for these two protein domains. This leads to the attractive hypothesis that the *spoIIG* gene product may act as a specific sigma-like factor and that the region it has in common with the genuine *E. coli* sigma factor may similarly be involved in binding to the core RNA polymerase. This hypothesis is strengthened by the known functional and structural relationship between the *E. coli* and *B. subtilis* core RNA polymerases^{19–21}. Moreover, recent nucleotide sequence analysis of the *B. subtilis* σ^{55} structural gene (R. H. Doi, personal communication) indicates that the homologous region in the *E. coli* sigma factor and the *B. subtilis spoIIG* gene product is conserved in the major sigma factor of *B. subtilis*, σ^{55} , confirming previous immunological experiments²². One obvious possibility is that the *spoIIG* gene encodes σ^{29} , a polypeptide associated with RNA polymerase purified from *B. subtilis* sporulating cells and conferring on it new promoter recognition specificity²³. σ^{29} is a M_r 27,000–29,000 polypeptide^{4,23,24} and appears about 1 h after the onset of sporulation²³, the exact moment at which the temperature-sensitive period starts in strain 279.1, a strain mutated in the *spoIIG* gene⁸. However, sporulation involves the expression of multiple gene sets and could rely on the presence of several alternative sigma-like proteins. The *spoIIG* gene product could equally be one of these, as yet uncharacterized, transcriptional factors.

We thank Françoise de la Torre for technical assistance in cloning the *spoIIG* gene, Isabelle Giri for help with computer sequence analysis and Jocelyne Mauger for typing the manuscript. This work was supported by a CNRS grant (LA 136) to P.S. and J.B.

Note added in proof: The region of homology between the *B. subtilis spoIIG* gene product and the *E. coli* sigma factor has recently been found to be conserved also in the heat shock regulatory protein of *E. coli*²⁷.

Received 18 June; accepted 6 September 1984.

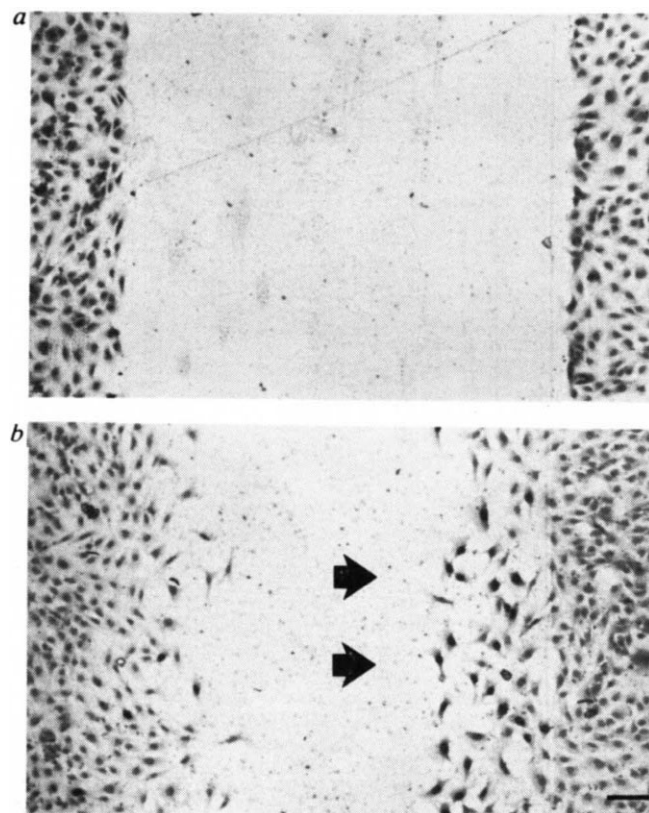
1. Piggot, P. J. & Coote, J. G. *Bact. Rev.* **40**, 908–962 (1976).
2. Henner, D. J. & Hoch, J. A. *Microbiol. Rev.* **44**, 57–82 (1980).
3. Losick, R. & Pero, J. *Cell* **25**, 582–584 (1981).
4. Fukuda, R. & Doi, R. H. *J. Bact.* **129**, 422–432 (1977).
5. Shimotsu, H., Kawamura, F., Kobayashi, Y. & Saito, H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 658–662 (1983).
6. Ramakrishna, N., Dubnau, E. & Smith, I. *Nucleic Acids Res.* **12**, 1779–1790 (1984).
7. Bouvier, J., Stragier, P., Bonamy, C. & Szulmajster, J. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
8. Young, M. J. *Bact.* **126**, 928–936 (1976).
9. Ayaki, H. & Kobayashi, Y. *J. Bact.* **158**, 507–512 (1984).
10. Bonamy, C. & Szulmajster, J. *Molec. gen. Genet.* **180**, 57–65 (1982).
11. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
12. McLaughlin, J. R., Murray, C. L. & Rabinowitz, J. C. *J. biol. Chem.* **256**, 11283–11291 (1981).
13. Burton, Z. *et al. Nucleic Acids Res.* **9**, 2889–2903 (1981).
14. Schwartz, R. M. & Dayhoff, M. O. *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 3 (ed. Dayhoff, M. O.) 353–358 (National Biomedical Research Foundation, Washington, D.C., 1978).
15. Needleman, S. B. & Wunch, C. D. *J. molec. Biol.* **48**, 443–453 (1970).
16. Smith, T. F., Waterman, M. S. & Fitch, W. M. *J. molec. Evol.* **18**, 38–46 (1981).
17. Nichols, B. P., Van Cleemput, M. & Yanosfsky, C. *J. molec. Biol.* **146**, 45–54 (1981).
18. Band, L., Shimotsu, H. & Henner, D. J. *Gene* **27**, 55–65 (1984).
19. Losick, R., Shorestein, R. G. & Sonenshein, A. L. *Nature* **227**, 910–913 (1970).
20. Shorestein, R. G. & Losick, R. *J. biol. Chem.* **248**, 6170–6173 (1973).
21. Achbeyer, E. C. & Whiteley, H. R. *J. biol. Chem.* **255**, 11957–11964 (1980).
22. Wong, S. L. & Doi, R. H. *J. biol. Chem.* **257**, 11932–11936 (1982).
23. Haldenwang, W. G., Lang, N. & Losick, R. *Cell* **23**, 615–624 (1981).
24. Linn, T., Greenleaf, A. L. & Losick, R. *J. biol. Chem.* **250**, 9256–9261 (1975).
25. Ehrlich, S. D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1433–1436 (1978).
26. Vieira, J. & Messing, J. *Gene* **19**, 259–268 (1982).
27. Landick, R. *et al. Cell* **38**, 175–182 (1984).

Erratum

New evidence that growth in 3T3 cell cultures is a diffusion-limited process

G. A. Dunn & G. W. Ireland
Nature **312**, 63–65 (1984)

BECAUSE of an editorial error, an incorrect figure was used in this letter. The correct Fig. 3 is shown below.



African marsupials— vicariance or dispersion?

UNTIL recently the marsupials were not recorded in Africa, but two recent discoveries in the Lower Eocene and in Lower Oligocene of North Africa have demonstrated the presence of this group of mammals on the continent. Their phylogeny and palaeobiogeography must, therefore, be reconstructed.

The discovery of marsupials in the Lower Oligocene of Africa has recently been interpreted as indicative of continuous land dispersion across the Tethys from Europe to Africa¹. Another recent discovery, however, proves that marsupials were present in northern Africa in the late Lower Eocene². The upper molar described from the north-west Sahara displays both primitive characters (the narrowness of the protocone) and very derived characters (dilambdodonty), which could suggest either immigration from a European centre of origin or survival of an ancestral stock common to South America and Africa before the opening of the South Atlantic (vicariance).

A vicariant origin for the African marsupials should not be excluded, because the time when the main marsupial dental characters appeared³, which would permit an evolutionary evaluation of African Oligocene and European late Eocene marsupials, is unknown. Unfortunately there is no way of testing the phyletic relationships between the Saharan and Egyptian forms, due to the scanty fossil record. On the other hand, European or Holarctic origin is supported by the continuous record of mammals of Holarctic affinities discovered in the Palaeocene⁴, Eocene^{5,6} and Oligocene⁶ of Africa, and by the absence of marsupials from the only known Palaeocene mammal fauna of Morocco⁴. The recent discovery of marsupials in the European Palaeocene⁷ reinforces this interpretation. However, several other palaeobiogeographic scenarios might also help to explain their presence in Africa. According to Sigé⁸, Africa could have been a land route between South America and North America, but there is no evidence of non-marsupial vertebrate dispersion to support this hypothesis.

Several authors have claimed that Africa could have been a link during early Cenozoic between South America and Europe for phorusrhacid birds⁹, mesosuchian ziphodont crocodiles¹⁰, characid fishes⁴ and marsupials⁷. As in the case of Africa, models which exclude Asia and the Indian plate need to be reconsidered. No true marsupial has been discovered in the Asian Mesozoic, but very few micromammals have been described from the rich mammalian assemblages of the Chinese early Tertiary. Additional information is provided by molecular data. For example, the structure and

sequence of neuropeptides¹¹ have recently shown that the most primitive living marsupials are Australian, thereby suggesting that Australia could have been the dispersion centre of the group. In conclusion, these new data demonstrate that the phylogeny and palaeobiogeography of marsupials is much more complicated than previously assumed, and should stimulate new research in many fields.

J. J. JAEGER
M. MARTIN

*Laboratoire de Paléontologie des
Vertébrés et de Paléontologie humaine,
CNRS ERA, 963, tour 25-15,
Université P. et M. Curie (Paris VI),
4 place Jussieu, 75230 Paris,
Cedex 05, France*

1. Bown, T. M. & Simons, E. L. *Nature* **308**, 447-449 (1984).
2. Mahboubi, M., Ameur, R., Crochet, J. Y. & Jaeger, J. J. *C. r. heb. Séanc. Acad. Sci., Paris D297*, 691-694 (1983).
3. Crochet, J. Y. *Les Marsupiaux du Tertiaire d'Europe*, 1-287 (Singer Polignac, Paris, 1983).
4. Cappetta, H. et al. *Géobios* **11**, 257-263 (1978).
5. Mahboubi, M., Ameur, R., Crochet, J. Y. & Jaeger, J. J. *Géobios* (in the press).
6. Simons, E. L. *Peabody Mus. nat. Hist.* **28**, 1-22 (1968).
7. Crochet, J. Y. & Sigé, B. *Palaeovertebrata* **13**, 57-64 (1983).
8. Sigé, B. *1ère réunion annuelle des Sciences de la Terre*, 382 (Soc. géol. Fr, 1973).
9. Mourer-Chauviré, C. *Géobios* **14**, 637-646 (1981).
10. Buffetaut, E., *Nature* **300**, 176-178 (1982).
11. Acher, R. *Trends Neuro. Sci.* **4**, 226-230 (1981).

BOWN AND SIMONS REPLY—Jaeger and Martin suggest a possible vicariant origin for newly discovered marsupials from the Lower Eocene of Algeria and the Oligocene of Egypt^{1,2}. These are the only metatherian remains ever recovered from the African continent, and the vicariance hypothesis suggests that these animals are survivors of an ancestral marsupial stock common to South America and Africa before the opening of the South Atlantic (before the early Cretaceous³). The distributions of at least the early Tertiary marsupials of Europe and North America might then have been derived from this remnant African stock, rather than having originated in Europe, as we have claimed² for the Egyptian marsupial, or in North America, as suggested by others^{4,5} for the Euramerican marsupials. Although a vicariance theory of origin of mammalian groups is in vogue, its application to this problem and to others is generally supported only by negative evidence, that is, the absence of a good fossil record. For African marsupials, a vicariant origin is the least parsimonious alternative and requires detailed speculation about the composition of a non-existent fossil record.

First, there are no known pre-late Cretaceous marsupials from South America (the late Cretaceous marsupials from Peru are virtually identical to North American late Cretaceous *Alphadon*⁶). Second, the only proposed marsupial of early Cretaceous antiquity from anywhere in the world is North American *Holo-*

*clemensia texana*⁷, a form that some feel is of doubtful metatherian affinity⁸⁻¹⁰. Even if *Holoclemensia* is metatherian, vicariance biogeography would place the origin of the Metatheria in North America, in accordance with most current thought. Third, the combination of primitive and derived characters in the Eocene Algerian form is unknown in any other didelphid, including those from South America, and the Egyptian marsupial shares derived characters with late Eocene and early Oligocene marsupials from Europe. By a vicariant origin, the Didelphidae of South America and those of Euramerica would have to be polyphyletic, or else the South America didelphids were at some time reintroduced there from Africa via Euramerica. Fourth, Egyptian *Peratherium* is so like European species of the genus, that had the Old World didelphids originated in Africa and spread to Euramerica, the avenue of dispersal between Africa and Europe must have been irregularly maintained over many millions of years. The first three points are not supported by palaeontological evidence, and the last observation is directly opposed to what is known of Tethyan palaeogeography during the early Tertiary.

Although Jaeger and Martin are correct in observing that a vicariance origin for the African marsupials should be examined, there is no evidence to support their hypothesis involving the composition and dispersion of an unknown fauna. We believe that there is some evidence that the Egyptian marsupial was an early Tertiary immigrant from Europe. Marsupials were probably present in Europe during the Palaeocene¹¹, although they are as yet unknown from mammalian microfaunas of that age in Africa. The Egyptian marsupial cannot have arrived in Egypt before the latest Eocene because of marine conditions there. Its appearance in the Egyptian continental Oligocene approximately coincides with the Tethyan regression, when Europe and Africa were presumably separated by less water distance than for many millions of years previously.

We also observe that the presence of a single stylar cusp on a narrow stylar shelf, in conjunction with the more or less dilambodont molar structure in the specimen from the Eocene of Algeria, is so unusual for a marsupial as seriously to jeopardize its assignment to the Metatheria. The tooth appears to share at least as many derived characters with European *Remiculus* (Mixodectidae) and *Adapisoriculus* (Insectivora, *incertae sedis*) as it does with any marsupial.

Finally, we did not "... interpret the discovery of marsupials in the Lower Oligocene of Africa as indicative of continuous land dispersion across the Tethys from Europe to Africa", as Jaeger and Martin suggest. Rather, we observed that the palaeogeographical reconstructions do

not favour a direct land connection of Africa with Europe at that time².

T. M. BOWN
E. L. SIMONS

Geological Survey,
Denver, Colorado 80225, USA
and
Duke University Primate Center,
Durham, North Carolina 27705, USA

1. Mahboubi, M., Ameur, R., Crochet, J.-Y. & Jaeger, J. J. *C. r. hebd. Séanc. Acad. Sci., Paris D297*, 691–694 (1983).
2. Bown, T. M. & Simons, E. L. *Nature* **308**, 447–449 (1984).
3. Smith, A. G. & Briden, J. C. *Mesozoic and Cenozoic Paleogeographical Maps* (Cambridge University Press, 1977).
4. Clemens, W. A. *Evolution* **22**, 1–18 (1968).
5. Crochet, J.-Y. *Geobios Mem.* 1, spéc., 127–134 (1977).
6. Sigé, B. C. *r. hebd. Acad. Séanc. Sci., Paris* **273**, 2479–2481 (1971).
7. Slaughter, B. H. *J. Linn. Soc. (Zool.)* **50**, 131–143 (1971).
8. Tedford, R. H. *Soc. Econ. Paleont. Miner. Spec. Publ.* **21**, 109–126 (1974).
9. Kielan-Jaworowska, Z., Eaton, J. G. & Bown, T. M. in *Mesozoic Mammals: The First Two-thirds of Mammalian History* (eds Lillegraven, J. A., Kielan-Jaworowska, Z. & Clemens, W. A.) (University of California Press, Berkeley, 1979).
10. Clemens, W. A. *Mesozoic Mammals: The First Two-thirds of Mammalian History*, Ch. 11 (University of California Press, Berkeley, 1979).
11. Crochet, J.-Y. & Sigé, B. *Palaeovertebrata* **13**, 57–64 (1983).

Cometary showers and unseen solar companions

THE existence of an unseen solar companion has occasionally been invoked to account for some unsolved astrophysical problem. The latest hypothesis, as proposed by Whitmire and Jackson¹ and by Davis *et al.*², is that such a companion, in a distant eccentric orbit, periodically passes through the Oort cometary cloud and causes a shower of comets from an also unseen inner Oort cloud reservoir. Impacts on the Earth from these repeated showers then results in periodic biological extinction events.

A shower of comets from the inner Oort cloud is expected to involve some 2×10^9 comets passing four times each within the orbit of the Earth, over a period of $\sim 10^6$ yr. This episodic flux is equivalent to the presently estimated steady-state cometary flux of 16 comets $\text{AU}^{-1} \text{yr}^{-1}$ (ref. 3) (brighter than $H_{10} = 11$) for a period of 500 Myr. Thus, the seven to nine cometary showers, spaced at 28-Myr intervals, for which a cratering record has allegedly been found⁴, would produce as many craters as the steady-state cometary flux over the Solar System's entire history.

Present estimates of the cratering rates on the Earth and Moon based on dated surfaces of very different ages⁵, are in agreement with one another, after allowing for differences in gravity-scaling and impact cross-section. Estimates of the expected cratering from Earth-crossing asteroids are roughly twice these values⁶; estimates of the cratering from comets range between 0.3 and 1.0 times the rate based on dated crater surfaces^{7,8}. Although it seems that the known Earth-crossing objects yield a total cratering rate on the Earth and Moon which is too high, the problem is not serious because of the large error bars in these calculations.

However, adding the flux predicted for cometary showers every 28 Myr makes the problem significantly worse. These showers result in an 18-fold increase in the mean cometary flux over the past 400 Myr, resulting in between 5.4 and 18 times as many additional craters on the Earth and Moon in that period, outside any reasonable error limits on these rates.

Even allowing this discrepancy, the probability that such a cometary shower would result in the impact of a 10-km diameter cometary nucleus on the Earth is about 0.055 per shower. Thus, major events like the Cretaceous-Tertiary (K-T) extinction might be expected on average every 510 Myr, considerably longer than is observed.

One can also consider the stability of the hypothesized 'death star' orbit. With a perihelion of 3×10^4 AU and a period of 28 Myr, the proposed companion star would have an aphelion of 1.54×10^5 AU. Monte Carlo simulation modelling of the dynamical evolution of a large sample of hypothetical stars starting with that orbit⁹ has shown that 23% escape the Solar System in 10 orbits or less, and 86% escape in $< 10^9$ yr. The average change in orbital period per orbit is 10%. Thus, it is unlikely that an unseen companion star could have remained in a constant period orbit, over the period suggested by the terrestrial extinction record, and impossible that it has been in that orbit since the origin of the Solar System.

If a more tightly bound original orbit were to be assumed, the problem of the increase in the terrestrial cratering record would only worsen because of the increased frequency of cometary showers implied. Also, the Oort cloud population would have been severely depleted by the repeated perturbations of the companion star, implying an immense initial cloud mass. If, on the other hand, one assumed that the companion star had been recently captured, an event with a probability of 10^{-13} , we would still expect an approximate doubling of the cometary cratering on the Earth in the past 300 Myr.

Finally, an additional problem with the death star scenario is that there do not seem to be many random events mixed with the periodic signal from cometary showers. Random Apollo asteroid impacts should produce a major extinction every 10^8 yr or so, and large (though not necessarily catastrophic) cometary impacts should occur with about one-third that frequency.

The proposed death star scenario, however interesting, has consequences which its proposers have failed to consider, and for which we have little evidence. If the periodicity in the fossil extinction record is indeed real, then some other mechanism must be found for creating it. It would seem wise to look for a period of which we are already aware, for example, the time between the Sun's galactic plane crossings is close to the required

value, though the extinction mechanism remains a mystery. Even so, this would seem a more worthwhile avenue for future study than continuing to chase an invisible star which does not exist.

PAUL R. WEISSMAN
Earth and Space Sciences Division,
Jet Propulsion Laboratory,
Pasadena, California 91109, USA

1. Whitmire, D. P. & Jackson, A. A. *Nature* **308**, 713–715 (1984).
2. Davis, M., Hut, P. & Muller, R. A. *Nature* **308**, 715–717 (1984).
3. Everhart, E. *Astr. J.* **72**, 1002–1011 (1967).
4. Alvarez, W. A. & Muller, R. A. *Nature* **308**, 718–720 (1984).
5. Grieve, R. A. F. & Dence, M. R. *Icarus* **38**, 230–242 (1979).
6. Shoemaker, E. M. in *Impact and Explosion Cratering* (eds Roddy, D. J., Pepin, R. O. & Merrill, R. B.) 617–628 (Pergamon, New York, 1977).
7. Weissman, P. R. *Geol. Soc. Am. spec. Pap.* **190**, 15–24 (1982).
8. Shoemaker, E. M. & Wolfe, R. F. in *Satellites of Jupiter* (ed. Morrison, D.) 277–339 (University of Arizona Press, 1982).
9. Weissman, P. R. *IAU Colloq.* **83** (in the press).

MULLER *ET AL.* REPLY—Weissman's objections are based, we believe, on a misunderstanding of the companion star model^{1,2} and a misinterpretation of the periodicity discovered in the crater data³. We will review the salient features of these before addressing Weissman's points directly.

In the companion star model, the Sun's companion has an orbit with a 26-Myr period, matching that seen in the mass extinctions⁴. During perihelion, it passes close enough to the Oort cloud of comets to precipitate a shower in the normally depleted 'loss cone' region swept clean by Jupiter and Saturn; this requires the companion's orbit to have an eccentricity greater than ~ 0.6 . The number of comets in the shower depends on the number in the inner part of the Oort cloud, but this number is not known within a factor of 10. If we take the nominal value of 10^{13} from the extrapolation of Hills⁵, we find the number of comets in a shower to be 2×10^9 and the number of impacts per shower to be 25; but these numbers could be as small as 4×10^8 and 5, respectively.

As stated in our original paper¹, and subsequently checked with detailed calculations⁶, the orbit we postulate is stable, with a half life of $\sim 10^9$ yr, not long enough for the companion star to have been in such a large orbit at the formation of the Solar System. As it is extremely unlikely that a star could have been captured by the Sun, the most likely scenario has the companion in a relatively tight orbit 5×10^9 yr ago, with perturbations from passing stars and molecular clouds over 5×10^9 yr expanding the orbit to the present one.

In Grieve's survey of impact craters on the Earth⁷, there are 58 craters with ages < 250 Myr, but only 29 of these have uncertainties in their ages of 20 Myr or less, and only 13 of these remain when we exclude the very recent (< 5 Myr old) craters; these appear in a table in our paper³. Our analysis of these craters showed a statistically significant periodicity, with period and phase identical

(within errors) to those found in the fossil record. On closer examination we found that most of the signal was coming from the 11 largest craters, those with a diameter >10 km. One would expect comets in a shower to be larger than the background comets if, for example, the comets from the inner cloud were more massive than those from the outer regions, as might be expected from an accretion model of comet formation. Of the 11 large craters, one is clearly not associated with the periodicity, and as many as two others may be chance coincidences. Thus, the evidence indicates that about 9 ± 3 of the 13 craters with precise ages originated from the periodic showers. (The uncertainty estimate includes both systematic and Poisson errors.) Stated another way, we conclude that $70 \pm 25\%$ of the craters were formed during showers.

Weissman claims that the total number of comets hitting the Earth from showers is substantially greater than the number hitting in the time between showers—this is not necessarily true in our model, and the impact data indicate that the numbers may be roughly equal. Perhaps Weissman's confusion comes from a poorly worded statement in our original paper¹ that the number of comets that arrive in a single shower "may be as much as one or two orders of magnitude greater than the number that arrive between showers". In this statement we were referring to the effects of the companion star alone, and were not including comets whose infall is triggered by random passing stars. Weissman states near the end of his letter: "there do not appear to be many random events mixed with the periodic signal"; as stated above, we believe that $30 \pm 25\%$ of the impact craters (principally those with diameters <10 km) may be random in arrival times.

Weissman says that the probability of a large comet (with a 10 km nucleus) hitting is small, and that such impacts are likely to occur every 500 Myr, a period "considerably longer than is observed". Only one such large impact has been proven in the last 250 Myr—the one at the end of the Cretaceous. The crater for this impact, if it still exists, should be 100–200 km across. (If the crater is on the sea floor it would not have been found.) All the other extinctions during this period were smaller, with associated craters 10–100 km across. Weissman claims there should be one large impact in 500 Myr; the data show one in 250 Myr. Although we do not necessarily accept Weissman's number (he does not give a derivation) we find no significant discrepancy with the data.

If the orbit of the companion star were once much closer to the Sun, as is necessary in our model, then Weissman objects that the showers would have been more intense in the past. This is true, but is not an objection. In fact, there is clear evidence that bombardment in the remote past was once much greater than in the

last 250 Myr. The 'late heavy bombardment' of the Moon (when most of the lunar craters were formed) ended 3.9×10^9 yr ago, conceivably when the companion star was scattered to a larger orbit. Weissman correctly states that an ancient close-in orbit would have severely depleted the Oort comet cloud, implying that there was originally an 'immense cloud mass'. As the mass of this early cloud could still have been much less than the mass of the companion star, we can see no valid objection. If the companion star exists, then we certainly will have to change our models of the dynamics of the early Solar System. But inconsistency with the models is not the same as inconsistency with the data on which those models were based. We know of no data inconsistent with our model of a solar companion star, and we know of no alternative model consistent with the measured period and phase of the impacts.

RICHARD A. MULLER*
PIET HUT†
MARC DAVIS‡
WALTER ALVAREZ§

* *Department of Physics and Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720, USA*

† *Institute for Advanced Study, Princeton, New Jersey 08540, USA*

‡ *Department of Physics and Astronomy, University of California, Berkeley, California 94720, USA*

§ *Department of Geology and Geophysics, University of California, Berkeley, California 94720, USA*

1. Davis, M., Hut, P. & Muller, R. A. *Nature* **308**, 715–717 (1984).
2. Whitmire, D. & Jackson, A. *Nature* **308**, 713–715 (1984).
3. Alvarez, W. & Muller, R. A. *Nature* **308**, 718–720 (1984).
4. Raup, D. & Sepkoski, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 801–805 (1984).
5. Hills, J. G. *Astr. J.* **86**, 1730–1740 (1981).
6. Hut, P. *Nature* **311**, 636–638 (1984).
7. Grieve, R. *Geol. Soc. Am. spec. Pap.* **190**, 25–37 (1982).

WHITMIRE AND JACKSON REPLY—The contribution of shower comets to the mean cratering rate need only be comparable to the (assumed) steady-state random contribution to produce a strong signal in the Fourier spectrum of dated craters. The number of terrestrial impacts resulting from a shower of 2×10^9 comets, brighter than $H_{10} = 11$ and with perihelia ≤ 1 AU, is given by the product of (2×10^9 comets) and (4 perihelia passages per comet) and (2×10^{-9} impacts per perihelion passage), which is equal to 16 impacts. According to Weissman¹, a long period comet, brighter than $H_{10} = 11$, has a mass $\geq 5 \times 10^{14}$ g and produces a crater ≥ 20 km in diameter. The mean cratering rate, for craters ≥ 20 km, over the last 300 Myr is $2.35 (\pm 0.13) \times 10^{-14} \text{ km}^{-2} \text{ yr}^{-1}$, which corresponds to ~ 50 impacts of comets, brighter than $H_{10} = 11$, per 28 Myr. Thus the shower contribution to the mean cratering rate is less than the total rate and comparable to the mean background rate.

According to Wetherill and Shoemaker³, a random 10-km asteroid is

expected to impact on the Earth about every 50 Myr. The absence of any random extinction events⁴, and the evidence that major extinctions occur over intervals of ~ 1 Myr, both suggest that the primary extinction mechanism may be associated with the enhanced impact rate (≥ 30 times background) during a shower, rather than with a single catastrophic impact, although such an event should also occur occasionally.

The stability of the companion's present orbit towards stellar perturbations and the galactic tides has recently been investigated by Hut⁵ and Hills⁶. These results indicate that the orbital period should random-walk away by 10–20% over a 250-Myr interval, a result not obviously incompatible with the cratering and extinction data. The evolution of the companion from a tighter orbit⁷, and the associated cratering rates, remain to be investigated in detail.

Sepkoski and Raup⁴ have recently extended their original analysis and have further restricted the uncertainty in the extinction period. They find the period to be 26.2 ± 1 Myr and conclude that this is incompatible with the 33 ± 3 -Myr interval between the Sun's galactic plane crossings.

DANIEL P. WHITMIRE
*Department of Physics,
University of Southwestern, Louisiana,
P.O. Box 44210, Lafayette,
Louisiana 70504, USA*

ALBERT A. JACKSON IV
*Computer Science Corporation,
1300 Bay Area,
Houston, Texas 77058, USA*

1. Weissman, P. R. *Geol. Soc. Am. spec. Pap.* **190**, 15–24 (1982).
2. Grieve, R. A. F. & Dence, M. R. *Icarus* **38**, 230–242 (1979).
3. Wetherill, G. W. & Shoemaker, E. M. *Geol. Soc. Am. spec. Pap.* **190**, 1–13 (1982).
4. Sepkoski, J. J. & Raup, D. M. *Dynamics of Extinction* (ed. Elliott, D.) (Wiley, New York, in the press).
5. Hut, P. *Nature* **311**, 638–641 (1984).
6. Hills, J. *Nature* **311**, 636–638 (1984).
7. Whitmire, D. P. & Jackson, A. A. *Nature* **308**, 713–715 (1984).

The maximum entropy method for data analysis

IN data analysis it is very satisfying when a useful technique, which has arisen in a more or less *ad hoc* manner, finds respectability as a manifestation of some formal theory. Skilling¹ invokes this argument for using the maximum entropy method for data analysis. Experience that the procedure often seems to work quite well is strengthened by reference to recent theoretical work. Skilling's conclusion is that maximum entropy is the only regularization method that should be used for a very wide range of data analysis. I suggest that this claim is unjustified on both pragmatic and theoretical grounds.

Let us confine our attention to one specific application, the enhancement of images as illustrated in Fig. 1 of Skilling's paper. What seems to me the major advantage of the maximum entropy method over many, but not all, other regularization

procedures, is the practical one that the resulting image, in being non-negative, automatically represents a valid solution. This results from the presence of $\log p_i$ in the formula for entropy. With other procedures, the values obtained for some p_i might be unrealistically negative. Even then, however, the complete solution could well be quite adequate in many applications and could well be attained with comparatively little computation.

In most applications of regularization, the regularization functional, exemplified in this method by entropy, S , incorporates some prior belief about the local smoothness of the true image. For example, suppose the pixels are ordered so that p_i and p_j are likely to be similar if i and j are close together; then one might consider, as an alternative to S , a criterion such as $-\sum(\log p_i - \log p_{i+1})^2$. At a higher level, one will know, *a priori*, that Fig. 1 is a smudged photograph of a car. The entropy functional does not permit representation of spatial ideas of this type. Certainly the cross entropy $S = -\sum p_i \log(p_i/m_i)$ allows for some acknowledgement of a prior model m but leaves the question of how to obtain m in any given application. There is the *ad hoc* suggestion of using the data to produce a preliminary m , but this would require using the data twice, first to find m and then to carry out the regularization. This is pragmatically reasonable, if not inevitable, but cannot be regarded as being formally respectable.

My principal point, however, is to dispute the main thrust of Skilling's article, that the maximum entropy method for data analysis, as used in the production of Fig. 1, stands in splendid isolation, on fundamental grounds. This statement is based on the axiomatic work of Shore and Johnson² and of Tikochinsky *et al.*³. They show that the choice of a probability distribution required to fit certain linear constraints has to be made on the basis of maximum entropy if certain axioms have to be satisfied. Typically, the effective constraints in the image-processing problems are highly nonlinear, and are expressed by the intensities themselves, not their proportions. It is not obvious that the mathematics of Shore and Johnson carry over to such a problem, and I cannot accept that their axioms, particularly those of subclass and system independence, are relevant to the regularization context. Tikochinsky *et al.*, in their discussion of reproducible experiments, specifically require the p_i to be probabilities; the assumption of linear constraints is vital, and even their definition of reproducible experiments does not correspond with the interpretation given by Skilling.

Thus, the maximum entropy method for data analysis is a useful tool, with some practical advantages and disadvantages over other regularization procedures. It is mildly interesting that there are links with the formal work on the principle of maximum entropy but the parallel simply

has not been shown to be exact enough to disallow flexibility in choosing a procedure to clarify Fig. 1 of ref. 1. As 'consistency', as defined by Shore and Johnson, seems irrelevant in this context, we need not fear the charge of inconsistency by rejecting entropy.

D. M. TITTERINGTON

Department of Statistics,
University of Glasgow,
Glasgow G12 8QW, UK

1. Skilling, J. *Nature* **309**, 748-749 (1984).
2. Shore, J. E. & Johnson, R. W. *IEEE Trans. Inf. Theory* **IT-26**, 26-37 (1980); **IT-29**, 942-943 (1983).
3. Tikochinsky, Y., Tishby, N. Z. & Levine, R. D. *Phys. Rev. Lett.* **52**, 1357-1360 (1984).

SKILLING REPLIES—I find the arguments for using maximum entropy to be deeply compelling. In fact, I am unable to improve on Titterington's own words: "the maximum entropy method ... stands in splendid isolation, on fundamental grounds". Quite so. However, he takes me to task for not going beyond the description of an image as a set of probabilities/proportions to incorporate "some prior belief about the local smoothness of the true image". In this he has raised a most important point: clearly, the simple entropy formula $S = -\sum p_i \log(p_i/m_i)$ does not include any such belief; just as clearly, maximum entropy is losing power by ignoring this, provided of course that it is realistic to expect smoothness, or spikiness, or cars, or whatever in one's images.

I hold that these more complicated situations should ideally be handled by extending a technique which is basically correct rather than by invoking *ad hoc* alternatives. As Titterington's preference for smoothness tends to zero, I would like his formulae to reduce to the simple entropy form which is the compelling solution to that problem. This is because the regularization function, which expresses my preference for different shapes of image, should be independent of the particular form of the data to be measured. Logically, my preferences precede the data, so that I should use the same function whether I have convolution data (as in the car example) or marginal data (as in the theoretical discussion).

Fortunately, maximum entropy can include such preferences in an easy and natural way. The trick is to develop the identification of p with the image. The simplest identification is to let i be a single index ranging over the cells of the image, and p_i the corresponding proportion of flux. However, we can also let i be a composite index ranging over pairs of cells, and p_i the corresponding product of fluxes (itself a proportion). The model m_i can now incorporate pair correlations between cells as well as simple position-dependent information. If we take i to be a highly composite index, the model can encode correspondingly subtle details of prior knowledge. The arguments for using maximum entropy still hold, and we have been exploring this development. True, I

do not (yet?) know how to encode 'smoothness' in an absolute as opposed to *ad hoc* fashion. Nevertheless, I hope I have answered constructively Titterington's first major comment.

Concerning the relevance of Shore and Johnson's axioms¹, I do not think that the distinction between linear and nonlinear constraints is important here. The form of the experimental constraints is quite separate from the form of one's prior belief about the observed image, as coded in the regularization formula. Shore and Johnson explicitly state that their mathematics include nonlinear constraints. They use the technical trick of replacing a (convex) nonlinear constraint by that set of linear inequality constraints which defines the hull of the nonlinear constraint. Tikochinsky *et al.*² do not state this, having had physical applications in mind, but clearly the same trick can be used.

If the experimental constraints include dimensional information on intensities, the maximum entropy image will be accompanied by a dimensional number describing its normalization. Defining the regularization in terms of proportions rather than intensities merely requires the resulting image to have the same shape whether the data refer to microwatts or megawatts. The point about the subclass and system independence axioms is not that they hold for an arbitrary collection of data (such as a blurred photograph), but that they are compelling for particular collections (such as marginals). If one particular type of data forces me to use entropy, then I will consistently use entropy for other types also.

I admit that the use of entropy in data analysis imposes an interpretative gloss on the Shore and Johnson¹ and Tikochinsky *et al.*² papers, but I was writing a News and Views article, and I believe that these authors themselves approve of the application. We have to do something in data analysis: I am prepared to use probabilistic results to help to find sets of proportions. Probabilities—not that we need them anyway—and proportions are isomorphic: probabilities are just proportions in a sample space, and to each set of proportions there corresponds a probability distribution. Finally, I return to Titterington's remarks defending flexibility of choice. Maximum entropy using the composite (multi-cell) identification allows great flexibility in choosing the analytical procedure, but it also gives a framework within which choices can be related quantitatively to specific types of prior knowledge and discussed in a logical as opposed to a pragmatic fashion.

JOHN SKILLING

Department of Applied Mathematics
and Theoretical Physics,
University of Cambridge,
Cambridge CB3 9EW, UK

1. Shore, J. E. & Johnson, R. W. *IEEE Trans. Inf. Theory* **IT-26**, 26-37 (1980); **IT-29**, 942-943 (1983).
2. Tikochinsky, Y., Tishby, N. Z. & Levine, R. D. *Phys. Rev. Lett.* **52**, 1357-1360 (1984).

Past probabilities

Patrick Rabbitt

Genius, Creativity & Leadership: Historiometric Inquiries.

By Dean Keith Simonton.

Harvard University Press: 1984. Pp.231. \$22, £17.60.

DEAN Simonton is a prescient man. While the rest of us glumly accept that our specialisms are of little interest to colleagues trapped in their own obsessions, he has developed a career as a "historiometrician". In any Common Room in the world, after the initial difficulty of explaining what this means, he can emerge as an expert on the Great Academic Post — Prandial Topics of All Time. Do higher degrees actually hinder creativity? Can we quantify charisma? Is history, especially the history of science, shaped by individuals or by Zeitgeists? Does the physical presence of a commander really affect the outcome of a battle? Are important scientific discoveries typically made independently by several people, or by isolated individuals? And much, much more.

Simonton offers more than dinner-table expertise. He is widely read, and has training in statistical methods the principles of which he tries to transmit in five appendices. Perhaps these will not help people without, or impress people with, a similar background, but he undertakes to bring recent analytical ideas to his material. This is interesting, since the *post hoc* detection of statistical associations in sub-sets from vast bodies of potential data can easily persuade us of very strange things.

Does Simonton's use of some up-to-date statistical techniques reduce uncertainties about assumptions of causal relationships in very large, uncontrollable, *post hoc* data samples? Well, not as such. When the logic behind selection of numbers for analysis is dubious, elaborate statistical models cannot help. For example, on the whole Simonton favours the view that higher education stunts intellectual flexibility. He re-analyses Cox's (1926) collection of biographical details on "301 geniuses" extant between 1450 and 1900. Very few of them took higher degrees — which, therefore, are at least uncorrelated with originality? The conclusion may well be correct, but Simonton's argument requires some odd assumptions.

One is that the content, and the social and career implications of "higher degrees" have not changed since 1450. Simonton apparently reassures himself by computing that the incidence of higher degrees *within* Cox's sample did not change between 1450 and 1900. This is interesting, but hardly germane. We can only discover whether incidence of higher degrees in the Cox sample is atypical by comparing it against incidence in the reference popula-

tions from which it was drawn. To do this we must take into account enormous population growth since 1450, the consequent, likely, changes in the *percentages* of the population who could achieve historical recognition of pre-eminence in successive centuries, and the changes in the percen-

IMAGE
UNAVAILABLE FOR
COPYRIGHT
REASONS

tages of the population examined who ever experienced higher (let alone "higher higher") education from 1450 to 1900. Even if such a tedious computation were attempted, the Cox sample of 301 spread over four-and-a-half rapidly changing centuries would certainly be too small to allow any definite conclusion. It is not that historiographers do their sums wrong, it is that apparently they would rather analyse inadequate data than hold their peace.

Historical evidence will always be patchy. This makes it important for historiographers to recognize when they do not have enough data to answer a question. Simonton is intrigued (pp. 165–166) by Stewart's analysis that the familial birth order of "British Prime Ministers" (how many examined?) and Presidents of the United States (38 studied) appears to correlate with the types of political situation faced by their cabinets. Elsewhere he convinces himself that birth order is a potent force in determining personal style, and

feels this shows that events call forth the necessary person. Perhaps, but hold on. We have at least five birth-orders (only child, eldest, youngest of two, middle of three or more and youngest of three or more) and if we take only the four political scenarios distinguished by Stewart (civil conflict; international crisis and warfare; peace; and revolution), to show that one of these variables (birth-order) predicts even 10% of the variance in assignment to the other (political scenario) would require data from a far larger sample than Stewart provides. There are, indeed, well-established associations between birth order and intelligence test scores (5% of variance or less) but samples of thousands have been necessary to detect them.

The book is shot through with intriguing examples. For instance, can we find out whether originality in musical composition is a function of historical period? Simonton considered the first six notes of 15,618 themes produced by 479 "classical" composers between 1500 (Josquin de Pres) and 1950 (Schoenberg and Shostakovich). He examined the first-order transitional probabilities between notes in each of these note strings and related them to a (here undefined) baseline computation for all transitional probabilities between notes. Note strings were classified as less original if their transition probabilities coincided with, and more original if they deviated from, the frequency of two-note transitional probabilities in the (undefined) source transition matrix. On this basis Simonton plots his composer-originality by year from 1500 to 1950 and finds, overall, a steady increase in thematic originality over time, with sinusoidal variations peaking at about 1650 (Palestrina) and 1920 (Schoenberg), and with a severe trough between 1750 and 1820 (Haydn, Mozart, Beethoven and Schubert).

I leave others to argue how far two-note transitional-probabilities can be a valid index of thematic originality because I am entranced by the possibility of a splendid computational tautology: if the reference transition probability matrix is, as one must suspect, computed from *the entire corpus of 15,618 themes collected over 450 years*, this would mean that the earlier a composer flourished the more chance his work had to influence his successors and so the greater was his likely *indirect* influence on the baseline against which he was judged. The steady increase in "originality" over 450 years is at once explained. Further, the more eminent a composer the more he would influence his contemporaries and successors, and so the more substantial would be his indirect contribution to the "baseline" matrix. Hence the historical plunge from "originality" shown by Mozart *et al.*, who are very like the matrix which their influence determined. If we leave out of consideration such perturbing factors as the increased diffusion of musical ideas between geo-

graphically remote cultures as communications improved over the centuries, it may be that we can extract one conclusion from Simonton's elaborate computation: those whose works have survived longest, and those whose works are most often performed, shape the "norms" in musical composition!

Sophisticated mathematical models may be misleading in another way — they may provide *precise descriptions of numerical data* which are quite implausible in terms of real-world *processes*. Simonton elegantly reviews data on changes in human creativity in politics, arts and sciences between the ages of 20 and 70. Alas, though cheerful exceptions abound, the unmistakable trend is for peaks to be reached between 35 and 45 with an accelerated decline thereafter. He neatly fits an algebraic model to the data (pp.110 and 111), and on p.109 describes the functional implications of this model:

Each creator uses up [*sic*] this supply of creative potential [determined individually — perhaps at birth?] during his or her productive career by transforming potential into contributions.

I wish that I could believe that, by prudent inertia, I may yet reserve some "creative potential", undissipated by book reviews and unclouded by time, to astound my clapped-out contemporaries when I reach my 80s.

Incidentally it is encouraging for common sense to note that Simonton's algebraic statement of his model does not actually have this, or any other, clear functional implication. It seems to be a neutral equation for fitting empirical data using one invariant constant (*e*), two other constants empirically derived for particular disciplines (*a* and *b*) and a fourth constant empirically derived for individuals (*c*). It is not a parsimonious equation.

Simonton is clearly a very hard-working academic, and seems sincerely anxious that his subject should help us to understand ourselves better, to learn to anticipate and to avoid political violence, and to give each other just credit for real achievements. His book will be marvellous bed-time reading for any academic with a curiosity about historical or statistical matters. Only part of its entertainment value lies in the detection of charming errors and non sequiturs. Simonton has an acute sense of *really* interesting, though possibly insoluble, problems, and whenever he seems to go wrong he sets us the important task of wondering whether, and precisely how, something useful may yet be said. In spite (surely?) of the most extreme temptations he brings to his material no detectable humour. But, for the reader with his own sense of fun, this will be a delightful book. □

Patrick Rabbitt is Professor of Gerontology and Cognitive Psychology at the University of Manchester.

Cue for cooperation

F.H. Bronson

Mammalian Semiochemistry: The Investigation of Chemical Signals Between Mammals.

By Eric S. Albone.

Wiley: 1984. Pp. 360. £29.50, \$57.

MOST mammals are nocturnal and by human standards they are quite small. Thus they rely heavily upon their chemical senses when interacting with other members of their own species, avoiding predators or searching for food. Chemical cues play many crucial roles even in mammals that are larger, more diurnal, and hence more visually-orientated. The term "semiochemical" encompasses all such cues, whether they originate from plants or animals, and incorporates the concept of communication via pheromones.

By its very nature, the broad study of chemical cueing requires close collaboration between the chemist and the biologist; as witnessed by its impact on entomology, the result can be immensely rewarding. Research on mammals, however, has thus far been dominated by behaviourists and physiologists interested strictly in communication between members of the same species. Dr Albone is one of the few chemists whose primary interest is in mammalian chemical cueing, and this is the first account of the subject written by a chemist. The strengths and weaknesses of the book are those that might be expected when a scientist of one discipline looks at a complex subject that has its roots in another, markedly different discipline.

The first two chapters of the book introduce the subject and note the potential contributions and limitations of chemical investigations; of particular interest here is the author's emphasis on the

probable complexity of most mammalian chemical cues. The next six chapters deal with the potential sources of the cues and the types of compounds that may arise from these sources. Thus the skin, scent glands, urine, breath and the reproductive tract are considered in detail. Unfortunately much of this discussion relies upon data collected in human beings, and in only a few cases are sources and compounds related directly to functions (this, it should be said, is not the fault of the author, but an accurate reflection of the state of our knowledge). In a further chapter that should be read by everyone interested in the topic, Dr Albone considers the important role played by micro-organisms in shaping odours. Finally, a guest contributor, Dr Stephen Shirley, has written a nice primer on the neurophysiology of the chemical senses.

While this is a historically important book, dealing as it does with a young and rapidly developing area, it will probably fail in its stated objective of bringing together biologists and chemists. Much of what the biologist finds exciting about the subject is treated only superficially — the integration of chemical cues with information provided by the mammal's other senses; the shaping of responses by experience and context; the adaptive functions of signalling systems; and so forth. On the other hand, one can hope that the book will attract more chemists into semiochemical research, and all biologists interested in chemical cueing certainly should read it because it will help them understand how a chemist approaches many of the problems on which they work. Ultimately, real progress will occur only through cooperation, and scientists in both disciplines need to understand much better how their opposite numbers view this challenging field. □

F. H. Bronson is Professor of Zoology at the University of Texas, Austin.

SOM introduction

P.W. Hawkes

Theory and Practice of Scanning Optical Microscopy.

By Tony Wilson and Colin Sheppard.

Academic: 1984. Pp.213. £25, \$39.50.

THERE are not many types of radiation for which high-quality, two-dimensional, intensity-sensitive recording media exist. Light and electrons are exceptional in this respect, so that (apart from a few early scanning electron instruments) scanning microscopes were developed long after their fixed-beam counterparts had become commonplace. It is, however, gradually being recognized that the optics of scanning systems and the sequential nature of the image-forming process possess

attractive features in their own right, with the result that commercial scanning electron microscopes have been with us for nearly 20 years and commercial scanning transmission electron microscopes for about a decade. Now, as this book convincingly demonstrates, the scanning optical microscope (SOM) is beginning to supplement its conventional ancestor.

No two people are better qualified to write a book on SOMs than Tony Wilson and Colin Sheppard, for almost all the work on the theory and performance of these instruments has come from their Department of Engineering Science in Oxford. In extremely simple language, and using mathematics that anyone familiar with Sections 9.5–10.5 of Born and Wolf's *Principles of Optics* will find easy to follow, they take us through the theory of image formation in the SOM. The two operating modes — essentially large

detector and small detector — are explored in great detail, though if I have a criticism of the book, it is here: the authors are so familiar with the difference between these modes (Type 1 and Type 2 or confocal microscopy) that they have not, I think, realized that this distinction may not seem so important to newcomers.

That apart, the book covers all of the principal aspects of scanning optical microscopy: the many detector modes and special operating conditions, depth discrimination and depth of field, ultra-high resolution, practical details of the optical elements of the SOM, applications to semiconductors and some newer non-linear techniques. There is also a chapter on a scanning microscope which, ironically, mimics the fixed-beam instrument by furnishing the signals from the individual pixels almost simultaneously but which re-

tains the attractive features of the confocal arrangement.

This is the first book to be devoted entirely to the SOM and it is remarkably successful. A vast amount of material is presented very readably within its 200 or so pages, and practising microscopists — many of whom do not yet know what scanning, particularly scanning combined with electronic or computer image enhancement, has to offer them — will find it very helpful. The analysis of image formation will also be of some interest to students of the STEM, though the analogy between the two families of instruments is richer and more complex than is indicated here. □

P.W. Hawkes is Maître de Recherches at the Laboratoire d'Optique Electronique du CNRS, Toulouse.

Flying tonight

P.A. Racey

Bats: A Natural History.

By John E. Hill and James D. Smith.
British Museum (Natural History)/University of Texas Press: 1984. Pp. 243. £15, \$24.95.

The Lives of Bats.

By Wilfried Schober.
Croom Helm/Arco: 1984. Pp. 200. £13.95, \$24.95.

The past decade has seen a burgeoning of research interest in the second largest order of mammals, the Chiroptera, or bats. Such research is now serviced by international conferences and symposia, as well as by several newsletters, one of which has recently evolved into a dedicated journal. There is a certain sense of urgency about current work, to discover more about these mammals as outstanding examples of evolutionary development and adaptation before they decline in numbers even further. Bats are under pressure throughout the world — the use of agricultural insecticides, loss of natural roosts in caves and deciduous trees, and more recently the increasing application of remedial timber treatments in houses used as roosts have all contributed to a drastic fall in populations.

Accompanying recent activity among scientists has been a similar rise in public enthusiasm for bats. So it is appropriate that two general books on the subject should now appear at the same time.

Hill and Smith's *Bats: A Natural History* contains a wealth of up-to-date information on the world's bats in a well-written and easily assimilated form. The publisher's blurb commends the book to both the amateur naturalist and the serious student of zoology. While the treatment somewhat varies in depth, my own feeling is that it will be particularly useful as an introductory account for professional

biologists as well as for students.

The book goes well beyond the natural history approach implied by the title, as the authors have taken every opportunity to include physiological information in their reviews of thermoregulation, reproduction, echolocation and flight. Surveys of food habits and feeding, and population ecology, are also included, together with a systematic treatment of all bats at the family level. Hill and Smith are both systematists, and in their essay on the origin and evolution of bats they put forward the controversial theory that the Megachiroptera and Microchiroptera are not closely related, and that wings and flight have developed independently in these two groups of flying mammals. A chapter on man and bats contains a fascinating section on folklore, superstitions and legends, and includes a great deal of hitherto unpublished material. By contrast the authors are at their least engaging when writing about form and function.

The text is illustrated with excellent line diagrams, although several of the photomontages are of poor quality and are not closely related to the text. But overall this is a very useful book, even if it lacks that unbridled enthusiasm for all aspects of bat biology and a concern about their plight that is so evident in Wilfried Schober's *The Lives of Bats*.

Schober is a neurobiologist who has studied bats all his professional life. His account of their general biology is as authoritative as it is readable, and is ideally suited to the increasing number of amateur bat enthusiasts looking for an introduction to the subject. It has been sympathetically translated from the original German by Sylvia Furness, and revised by Tony Hutson, and is well illustrated with colour and half-tone plates which Schober has collected from colleagues around the world. □

P.A. Racey is a Senior Lecturer in Zoology at the University of Aberdeen.

Molecular orbitals: into the age of computer graphics

Jeremy K. Burdett

A Pictorial Approach to Molecular Structure and Reactivity.

By Robert F. Hout Jr, William J. Pietro and Warren J. Hehre.
Wiley: 1984. Pp. 404. £46.20, \$99.95.

W.L. Jorgensen and L. Salem's classic work, *The Organic Chemist's Book of Orbitals*, was published by Academic Press in 1973. It is a mark of changing times that whereas that volume included only rather crude, wire-frame drawings of orbitals, the present one contains glossy, computer-generated, Daliesque visions of the molecular orbitals of simple systems.

Indeed this is quite an unusual book altogether. It is largely composed of photographs of orbitals, with a couple of dozen pages of supporting text. For those of us who construct molecular orbital diagrams by doodling on the back of envelopes, it is pleasing to see how the shapes of the orbitals we sketch really look, and that *ab initio* computations confirm the general results that we get from simple ideas. Some of the photographs in this enormous undertaking are somewhat murky, while others are very complicated and resemble a piece of small intestine. But overall the pictures are remarkably effective in depicting the shapes of orbitals as they really are. Surely the age of computer graphics has arrived!

For the chemist who has an understanding of how the level structures of these molecules arise, some interesting points arise in the book. The differences in the shapes of the orbitals of the isoelectronic sequence CH_3 , NH_3 , OH_3^+ and PH_3 , for example, are really very striking. The spectrum of chemical behaviour shown by these species is much easier to understand with the assistance of these pictures than with the information provided by a conventional energy-level diagram.

I have one main criticism of the book. Although there are indeed nuggets of wisdom in the text supporting the pictures, there is just not enough in the way of chemical comment. For example the orbitals of ferrocene (the *bête noire* of *ab initio* theorists) command six pages of pictures and a measly four lines of text. There is much more to be said of the chemistry associated with the pictures than the authors give us here. Although in this sense the book does not live up to its title, it is by no means an unworthy successor to Salem and Jorgensen. □

Jeremy K. Burdett is a Professor in the Department of Chemistry at the University of Chicago.

Developing theories of evolution

J.S. Jones

Mechanisms of Morphological Evolution: A Combined Genetic, Developmental and Ecological Approach.

By Wallace Arthur.

Wiley: 1984. Pp.275. £19, \$36.

ANCIENT philosophy had a convincing explanation for earthquakes: they result from uneasy movements by the tortoise upon which the globe rests. Biologists have long been satisfied by similar explanations of the enigmatic in terms of the mysterious. In the study of evolution such theories include the role of the devil in placing fossils in rocks, of "hopeful monsters" in the origin of species, and — most recently — of mutations in the genes which are supposed to control development in producing the transitions between major taxonomic groups.

The preface to Wallace Arthur's book tells us, once again, that "current evolutionary theory in terms of neo-Darwinism is inadequate as an explanation of long term evolution". The remaining 275 pages largely fail to provide a convincing alternative. Unlike many post-Darwinians, Arthur accepts that modern evolutionary theory is not necessarily wrong simply because it is widely accepted. Indeed, parts of his own theory have a refreshing conventionality as, in the model put forward here, natural selection plays a major role in large scale evolution. Selection is said to act on the "D-genes" which are arranged as a "morphogenetic tree" to control development. "N-selection" depends on competition among members of a species while "W-selection" acts on the course of gene frequency change in expanding populations.

The book is an ambitious attempt to synthesize information from ecology, developmental biology and genetics. It is relatively free of what was known to an earlier and perhaps more cynical generation of evolutionists as "Wadswallow" — the posy of analogies (epigenetic landscapes, the relation between genotype and phenotype seen as a circus tent) with which even as distinguished an embryologist as C.H. Waddington was forced to disguise our naked ignorance of the genetic basis of body form. However, it remains the case that the theory developed here rests heavily on a class of genes whose very nature is speculative, and on modes of selection which have never been demonstrated to occur in natural populations.

One of the strengths of the book lies in

the author's ability to draw on the literature of developmental biology as much as on that of evolution. But although the references cited include a number from 1983, it is only in the past few months that information has begun to appear which suggests that there is indeed an evolutionary relationship between the genes which control development in different animals. We now know that the genetics of early development of insects with body forms as different as grasshoppers and flies is extremely similar, and that there has been little change in some of the genes which underlie differentiation in organisms separated by many millions of years of evolution. For example, the homeo box is a sequence which is closely related to mutations causing developmental abnormalities in *Drosophila*. It is found also in *Xenopus*,

chickens and human beings, suggesting that there is a fundamental similarity in the processes of pattern formation which has been conserved in the face of great changes in form during evolution. These genes regulating development therefore appear to be more — rather than less — resistant to the accumulation of mutations than are many of those which code for structural proteins; an observation which contrasts with the prediction of theories which emphasize the importance of developmental mutants in the origin of major groups. In a few years we may find that Darwinism has taught us more about the genetics of development than developmental biology can tell us about the processes of evolution. □

J.S. Jones is in the Department of Genetics and Biometry, University College London.

In the soup

A. E. Walsby

The Ecology of Freshwater Phytoplankton.

By C. S. Reynolds.

Cambridge University Press: 1984.

Pp.384. Hbk £30, \$54.50; pbk £12.50, \$24.95.

EACH springtime in lakes of temperate climes brings the stirring of a thin alphabet soup of phytoplankton: *Asterionella*, the *Biddulphiales*, *Closterium*, *Dinobryon* and the *Euglenales* . . . The soup thickens as the temperature rises and the days lengthen; species *a* outcompetes *b* to form the spring bloom this year; species *c* which won the race last year when the weather was colder, has not shown itself at all; and noxious *d*, which dominated before the sewage outfall was diverted, has thankfully kept a low profile. Such are the interests of those that study the phytoplankton.

The general principles that govern phytoplankton growth — light needed for photosynthesis, nutrients for the accumulation of biomass — are well-known. What Reynolds's book brings us is a quantitative appreciation of the things that limnologists have been generalizing about for decades. Take, for example, the problem of a population of diatoms slowly sinking through an unstratified wind-stirred lake. Reynolds shows by simple algebraic equations that in less than five times the minimum settling period under still-water conditions, 99% of the cells in continuously-stirred water will have reached the bottom mud. This surprising conclusion holds irrespective of the degree of turbulence. Here and elsewhere in the book equations are clearly derived for the biologist who can understand elementary mathematics even though he may not be good at doing it. Reynolds works with the same clarity and precision through the quantitative analysis of many other

physical and biological phenomena, including the changing light field, photosynthesis and productivity, the kinetics of growth and nutrient uptake, density gradients and terminal sinking velocities. Even apparently subjective topics, such as algal succession and population diversity, are given a similar numerical treatment.

One result of this rigorous quantifying is another alphabet soup, a text full of symbols; nearly a hundred are brought to order in the glossary. The basic units of the SI system are used in (almost) all the quantities, so that it is easy to build from one equation to the next and to follow even those with many terms. I found only a couple of dimensional errors.

The other strength of this book is that it recognizes the importance of knowing the special story of each alga's life cycle and predilections. For example, dinoflagellates are strong swimmers and do well on calm days in stratified lakes, apparently selecting just the right depth for their needs. Diatoms, which sink out like bricks under these conditions, enjoy the turbulence of early spring when every phytoplankton cell drifts at the mercy of the currents. Within the diatoms small differences in size and shape will determine such unrelated features as whether a cell can be re-suspended from the sediment, and how fast it will take up a particular nutrient.

The book contains a thorough treatment of much of the current knowledge on phytoplankton and what causes them to wax and wane in lakes. It is not just a derivative account but abounds with original ideas. The content is precise, and even picky (at every mention the word "alga" is in quotation marks, apparently because of its loose taxonomic meaning), but the style of writing is almost conversational and the book may be used for a good read as well as for reference. □

A.E. Walsby is Melville Wills Professor of Botany at the University of Bristol.

● *Quaternary Extinctions*, reviewed in last week's *Nature* (p. 225), is distributed in Britain and the rest of Europe by Eurospan, 3 Henrietta St, London WC2E 8LU. Price £63.50.

Astronomy

Astronomy and Astrophysics Abstracts, Vol.34. S. BÖHME *et al.* (eds). Springer-Verlag: 1984. Pp.921. ISBN 3-540-13485-9/0-387-13485-9. DM 189, \$68.80.

Bau und Physik der Galaxis. By HELMUT SCHEFFLER and HANS ELSASSER. Bibliographisches Institut, Mannheim: 1984. Pp.642. ISBN 3-411-01641-8. Np.

The Case for Mars, Vol.57 Science and Technology Series. PENELOPE J. BOSTON (ed.). American Astronomical Soc.: 1984. Pp.322. Hbk ISBN 0-87703-197-5; pbk ISBN 0-87703-198-3. Hbk \$45; pbk \$35.

Guidance and Control 1984. Advances in the Astronautical Sciences, Vol.55. ROBERT D. CULP and PARKER S. STAFFORD (eds). American Astronomical Society: 1984. Pp.485. Hbk ISBN 0-87703-1991-1; pbk ISBN 0-87703-200-9. Hbk \$60; pbk \$50.

Planetary Interiors. By WILLIAM B. HUBBARD. Van Nostrand Reinhold: 1984. Pp.334. ISBN 0-442-23704-9. \$42.50.

Secrets of the Sun. By RONALD GIOVANELLI. Cambridge University Press: 1984. Pp.116. ISBN 0-521-25521-X. £11.95, \$19.95.

Space: A Developing Role for Europe, Vol.56 Science and Technology Series. L.J. CARTER and PETER M. BAINUM (eds). American Astronomical Society: 1984. Pp.268. Hbk ISBN 0-87703-193-2; pbk ISBN 0-87703-194-0. Hbk \$45; pbk \$35.

Star Formation Workshop, Edinburgh 1983. R.D. WOLSTENCROFT (ed.). Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ: 1984. Pp.227. Pbk ISBN 0-902553-30-5. Np.

Physics

Acoustical Imaging, Vol.13. M. KAVEH, R.K. MUELLER and J.F. GREENLEAF (eds). Plenum: 1984. Pp.605. ISBN 0-306-41717-0. \$89.50.

Advances in Atomic and Molecular Physics. DAVID BATES (ed.). Academic: 1983. Pp.465. ISBN 0-12-003819-6. \$85.

Advances in Nuclear Physics. J.W. NEGELE and ERICH VOGT (eds). Plenum: 1984. Pp.289. ISBN 0-306-41524-0. Np.

Coherence and Quantum Optics V. LEONARD MANDEL and EMIL WOLF (eds). Plenum: 1984. Pp.1260. ISBN 0-306-41517-8. \$150.

Compendium of Thermophysical Property Measurement Methods. Survey of Measurement Techniques, Vol.1. K.D. MAGLIC, A. CEZAIIRLIYAN and V.E. PELETSKY (eds). Plenum: 1984. Pp.789. ISBN 0-306-41424-4. \$97.50.

A Concept of Mathematical Physics: Models for Space-Time. By TAMÁS MATOLCSI. Akadémiai Kiadó, Budapest: 1984. Pp.236. ISBN 963-05-3245-X. £15.50.

Crystal Optics with Spatial Dispersion, and Excitons. Springer Series in Solid-State Sciences 42. By V.M. AGRANOVICH and V.L. GINZBURG. Springer-Verlag: 1984. Pp.441. ISBN 3-540-11520-X/0-487-11520-X. DM 132, \$51.80.

Diffusion: Mass Transfer in Fluid Systems. By E.L. CUSSLER. Cambridge University Press: 1984. Pp.525. ISBN 0-521-23171-X. £32.50, \$49.50.

The Fokker-Planck Equation: Methods of Solution and Applications. By H. RISKEN. Springer-Verlag: 1984. Pp.454. ISBN 3-540-13098-5/0-387-13098-5. DM 125, \$46.70.

Lecture Notes in Physics, 200. Die Physik der Zeitrichtung. By H.D. ZEH. Springer-Verlag: 1984. Pp.86. Pbk ISBN 3-540-13336-4/0-387-13336-4. DM 18, \$7.10.

Lecture Notes in Physics, 201. Group Theoretical Methods in Physics. G. DENARDO, G. GHIRARDI and T. WEBER (eds). Springer-Verlag: 1984. Pp.518. Pbk ISBN 3-540-13335-6/0-387-13335-6. Np.

Local Density Approximations in Quantum Chemistry and Solid State Physics. JENS PEDER DAHL and JOHN AVERY (eds). Plenum: 1984. Pp.851. ISBN 0-306-41667-0. Np.

Long-Wave Optics: The Science and Technology of Infrared and Near-millimetre Waves. Vol.1 Principles, Vol.2 Applications. By G. W. CHANTRY. Academic: 1984. Vol.1 pp.429, ISBN 0-12-168101-7. Vol.2 pp.790, ISBN 0-12-168102-5. Vol.1 \$60, £45. Vol.2 \$55, £40.

Magnetohydrodynamic Equilibrium and Stability of Stellarators. By FRANCES BAUER, OCRAVOI BETANCOURT and PAUL GARABEDIAN. Springer-Verlag: 1984. Pp.196. ISBN 0-387-90966-4/3-540-90966-4. DM 110, \$40.10.

Polarons and Excitons in Polar Semiconductors and Ionic Crystals. J.T. DEVREESE and F. PEETERS (eds). Plenum: 1984. Pp.471. ISBN 0-306-41498-8. \$72.50.

The Principles of Nonlinear Optics. By Y.R. SHEN. Wiley: 1984. Pp.563. ISBN 0-471-88998-9. £49.45.

Quantum Questions: Mystical Writings of the World's Great Physicists. KEN WILBER (ed.). Shambhala Publications: 1984. Pp.208. Pbk ISBN 0-394-72338-4. £8.95.

Quantum, Space and Time — The Quest Continues. ASIM O. BARUT, ALWYN VON DER MERWE and JEAN-PIERRE VIGIER (eds). Cambridge University Press: 1984. Pp.659. Pbk ISBN 0-521-31911-0. Pbk £25, \$49.50.

Quarks: The Stuff of Matter. By HARALD FRITZSCH. Penguin: 1984. Pp.232. Pbk ISBN 0-14-022470-X. Pbk £4.95.

Rare Gas Solids. Springer Tracts in Modern Physics, Vol.103. By H. COUFAL *et al.* Springer-Verlag: 1984. Pp.99. ISBN 3-540-13272-4/0-387-13272-4. DM 64, \$25.10.

Semiconductors and Semimetals, Vol.20. Semi-Insulating GaAs. Academic: 1984. Pp.379. ISBN 0-12-75120-8. \$77, £59.50.

Solid State Physics, Vol.38. HENRY EHRENREICH and DAVID TURNBULL (eds). Academic: 1984. Pp.390. ISBN 0-12-607738-X. \$58, £41.

Spectrometric Techniques, Vol.III. GEORGE A. VANASSE (ed.). Academic: 1984. Pp.334. ISBN 0-12-710403-8. \$53.

Two-Dimensional Systems, Heterostructures, and Superlattices. Springer Series in Solid-State Sciences 53. G. BAUER, F. KUCHAR and H. HEINRICH (eds). Springer-Verlag: 1984. Pp.293. ISBN 3-540-13584-7/0-387-13584-7. DM 78, \$28.40.

Vibrations and Waves in Physics, 2nd Edn. By IAIN G. MAIN. Cambridge University Press: 1984. Pp.356. Hbk ISBN 0-521-26124-4; pbk ISBN 0-521-27846-5. Hbk £35, \$69.50; pbk £9.95, \$19.95.

VLSI Electronics Microstructure Science. NORMAN G. EINSPRUCH (ed.). Academic: 1983. Pp.401. ISBN 0-12-234107-4. \$59.50.

VLSI Electronics Microstructure Science. Materials and Process Characterization. NORMAN G. EINSPRUCH and GRAYDON B. LARRABEE (eds). Academic: 1983. Pp.601. ISBN 0-12-234106-6. \$79.

XIV International Symposium on Multiparticle Dynamics 1983. P. YAGER and J.F. GUNION (eds). World Scientific: 1984. Pp.1193. ISBN 9971-966-41-7. Np.

Chemistry

Advanced Organic Chemistry, 2nd Edn. Part A: Structure and Mechanisms. By FRANCIS A. CAREY and RICHARD J. SUNDBERG. Plenum: 1984. Pp.726. Pbk ISBN 0-306-41198-9. \$59.50.

Advances in Heterocyclic Chemistry, Vol.34. A.R. KATRITZKY (ed.). Academic: 1983. Pp.450. ISBN 0-12-020634-X. \$89.00.

Advances in Inorganic Chemistry and Radiochemistry. H.J. EMELEUS and A.G. SHARPE (eds). Academic: 1983. Pp.370. ISBN 0-12-023626-5. \$64.50.

Bio-Inorganic Chemistry. By ROBERT W. HAY. Wiley: 1984. Pp.210. Pbk ISBN 0-470-20066-9. Np.

Electrochemical Cell Design. RALPH E. WHITE (ed.). Plenum: 1984. Pp.398. ISBN 0-306-41805-3. Np.

General Chemistry in the Laboratory. By JULIAN ROBERTS, J. LELAND HOLLENBERG and JAMES M. POSTMA. W.H. Freeman: 1984. Pp.464. Pbk ISBN 0-7167-1574-0. \$18.95.

Grenzflächen-aktive Äthylenoxide-Addukte Ihre Herstellung, Eigenschaften, Anwendung und Analyse. By NIKOLAUS SHONFELDT. Wissenschaftliche Verlagsgesellschaft mbH, Stuttgart: 1984. Pp.602. ISBN 3-8047-0735-1. DM 254.

Quantum Chemistry, 3rd Edn. By IRA N. LEVINE. Allyn & Bacon: 1983. Pp. 566. ISBN 0-205-07793-5. Np.

Earth Sciences

The Caribbean-South American Plate Boundary and Regional Tectonics. WILLIAM E. BONINI, ROBERT B. HARGRAVES and REGINALD SHAGAM (eds). Geological Society of America: 1984. Pp.421. ISBN 0-8137-1162-2. Np.

Coasts: An Introduction to Coastal Geomorphology, 3rd Edn. By ERIC C.F. BIRD. Basil Blackwell: 1984. Pp.320. Hbk ISBN 0-631-13567-7; pbk ISBN 0-631-13568-5. Hbk £19.50; pbk £7.50.

Complexation of Trace Metals in Natural Water. Developments in Biogeochemistry. C.J.M. DRAMER and J.C. DUINKER (eds). Martinus Nijhoff: 1984. Pp.472. ISBN 90-247-2973-4. \$67.50, £41.95.

Fluvial Hydrology. By S. LAWRENCE DINGMAN. W.H. Freeman: 1984. Pp.383. ISBN 0-7167-1452-3. \$31.95.

The Gulf Coast. By C.C. LOCKWOOD. Louisiana State University Press: 1984. Pp.150. ISBN 0-8071-1170-8. £28.45.

Landforms of New Zealand. J.M. SOONS and M.J. SELBY (eds). Longman: 1984. Pp.392. Pbk ISBN 0-582-71786-8. Pbk £10.95.

Progress in Oceanography, Vol. 12. MARTIN V. ANGEL and JAMES J. O'BRIEN (eds). Pergamon: 1984. Pp.473. ISBN 0-08-031504-6. £94, \$150.

Sea Cliffs, Beaches, and Coastal Valleys of San Diego County: Some Amazing Histories and Some Horrifying Implications. By GERALD G. KUHN and FRANCIS P. SHEPARD. University of California Press: 1984. Pp.193. ISBN 0-520-05118-1. \$22.50.

Strain Patterns in Rocks. S.H. TREAGUS (ed.). Pergamon: 1984. Pp.470. ISBN 0-08-030273-4. £24.50, \$44.

Biological Sciences

Acarology VI, Vol.1. D.A. GRIFFITHS and C.E. BOWMAN (eds). Wiley: 1984. Pp.645. ISBN 0-470-27410-7. £49.50.

Adrenal Diseases in Childhood. Pediatric and Adolescent Endocrinology Vol.13. M.I. NEW and L.S. LEVINE (eds). Karger: 1984. Pp.235. ISBN 3-8055-3777-8. DM 98.99.

Advances in Cell Culture, Vol.3. KARL MARAMOROSCH (ed.). Academic: 1984. Pp.319. ISBN 0-12-007903-8. \$55.

Advances in Protein Chemistry, Vol.36. C.B. ANFINSEN, JOHN T. EDSALL and FREDERIC M. RICHARDS (eds). Academic: 1984. Pp.395. ISBN 0-12-034236-7. \$52, £36.50.

Algeny. A New Word — A New World. By JEREMY RIFKIN. Penguin: 1984. Pp.298. Pbk ISBN 0-14-007106-7. Pbk £2.95, \$6.95.

Animal Life-Cycle Feeding and Nutrition. By TILDEN WAYNE PERRY. Academic: 1984. Pp.319. ISBN 0-12-552060-3. \$24.50, £17.50.

Asian Marine Biology 1. THE MARINE BIOLOGICAL ASSOCIATION OF HONG KONG. Hong Kong University Press: 1984. Pp.175. Pbk ISBN 962-209-113-X. \$20.

Atlas of Skin Repair. By JANOS ZOLTAN. Karger: 1984. Pp.304. ISBN 3-8055-3511-2. DM 192, \$96.

The Audubon Society Handbook for Butterfly Watchers. By ROBERT MICHAEL PYLE. Charles Scribner: 1984. Pp.274. ISBN 0-684-18151-7. \$17.95.

Biology and Chemistry of the Carbonic Anhydrases. Annals of the New York Academy of Sciences, Vol.429. RICHARD E. TASHIAN and DAVID HEWETT-EMMETT (eds). New York Academy of Sciences: 1984. Pp.640. Hbk ISBN 0-89766-252-0; pbk ISBN 0-89766-253-9. Pbk \$140.

Biology of the Eucestoda, Vol.1. C. ARME and P.W. PAPPAS (eds). Academic: 1984. Pp.296. ISBN 0-12-062101-0. \$55.00, £34.00.

The Biology of Mistletoes. MALCOLM CALDER and PETER BERNHARDT (eds). Academic: 1984. Pp.348. ISBN 0-12-155055-9. \$49, £35.

Brain and Information: Event-Related Potentials. R. KARRER, J. COHEN and P. TUETING (eds). New York Academy of Sciences: 1984. Pp.768. Pbk ISBN 0-89766-242-3; hbk ISBN 0-89766-243-1. Pbk \$169.00.

Cardiovascular Physiology. By HENRY S. BADEER. Karger: 1984. Pp.276. ISBN 3-8055-3796-4. DM 69, \$34.75.

New skills for new technologies

from Richard Pearson

In some areas — biotechnology being an exception — employers requiring graduate skills seem reluctant to take on candidates with higher degrees.

THE United Kingdom, like many other countries, is living with a shortage of experienced information technology specialists while demand for these skills continues to grow. Whereas higher education has provided the basic skills and training for most of those entering the new technologies, the role of employers in training and developing new skills is now being emphasized. A UK government committee on the shortage of information technology skills has highlighted the need for a new "partnership between industry and higher education", and the University Grants Committee in its policy proposals for the 1990s stresses the key role to be played by "market signals" and the need for industry to improve its image if suitable students are to be attracted.

What information is available about the pattern of skills needed for the new technologies to help industry, education and students assess their training priorities? The United States has a vast array of studies and data about occupational trends, both past and future, as illustrated in last month's Employment column. In the United Kingdom, however, the only national forecast available does not provide anything like the necessary level of disaggregation: for example, it combines all scientists and engineers into a single category. Reliance therefore has to be placed on a series of *ad hoc* and often independently generated studies. What, then, do they reveal about the skills needed for the new technologies?

The term "new technologies" is, of course, a label of convenience and many of the constituent technical advances have largely been derived from existing technologies. Thus microelectronics is the miniaturization of electronic components, giving rise to major reductions in size, weight, power consumption and cost — the advantages which allow major new applications. Similarly, biotechnology has been used for centuries in brewing and fermentation, although it is the advances in genetic manipulation and bioprocess technology that have shown its potential in products as diverse as pharmaceuticals and feedstocks. Thus, although these technologies are "new", close examination shows that advances tend to be incremental; their contribution to society is determined as much by their application as by technological breakthroughs. The new technologies are found not only in the "sunrise" industries — they also have an important role to play in

revitalizing the "sunset" industries.

Employers' recruitment of staff to apply the new technologies similarly focuses largely on extensions of existing disciplines and emphasizes the key role of undergraduate degree programmes. The electronics industry relies almost exclusively on graduate intake to meet its future needs for professional staff, — the traditional entry routes, of promotion from craft or technician ranks, and part-time study having almost disappeared. Most employers are seeking graduates with broad-based electronics or related skills at a first degree level — even in highly specialized areas, such as semiconductor design and production, the preference is still for a majority of recruits from good quality first degree programmes rather than postgraduate courses.

The picture that emerges for information technology and computing is similar, although a large proportion of the intake to computing functions is of graduates of any discipline. In robotics, the demand is for graduates who can link mechanical, electronic and software skills, and not for fundamental science or a highly specialized degree.

Of all the new technologies, it is in biotechnology that postgraduate qualifications are most relevant. Again the disciplines sought are not new, but built around existing programmes. Here industry is much more closely linked to basic science, and few professional staff in biotechnology research have less than a PhD and, say, three further years of research experience.

Thus although the new technologies have a high dependency on higher education for their supply of skilled manpower, few of the skills initially sought are highly specific or dedicated to one type of narrow technology, discipline or employer. Rather, they allow for some flexibility in terms of course content and indeed a degree of substitution, for example, between applied physics and electronics. A further advantage is that these demands are also not numerically large when compared with an annual graduate output of around 100,000 per year — electronics and information technology each taking several thousand new graduates each year and biotechnology taking only a few hundred. These needs, however, are not being met.

The pattern and extent of shortages, as measured by vacancies that employers are unable to fill is, of course, greatly

influenced by past supply, present need and, particularly, the state of the economy. Thus the serious shortages of electronics and computing skills of 1978–80 abated during the recession but began to reappear in 1984. This year, the main areas of skill shortages have largely been restricted to electronics and computing staff with 3–7 years of experience, together with numerically smaller shortages in areas such as robotics, control, and telecommunications engineering and bioprocess engineering. There is, however, evidence that the incipient shortages of newly qualifying electronics and computing science graduates will soon increase. There are also many areas of work where existing specialists do not have the appropriate personal qualities or motivation to move into line management areas or the application of new technology, where there is a major need for technical skills.

Several initiatives to improve future supply are under way. The number of places on first degree courses in information technology is being increased for the second half of the 1980s while conversion courses are helping to alleviate short-term difficulties (see *Nature* 309, p100; 1984). Industry is also becoming increasingly involved with higher education through the provision of staff and equipment, and in course planning. A broader cross-section of students, however, must be encouraged to apply for the growing number of places, and especially those with the potential to apply such skills in line management and non-technical areas. Women represent a major untapped source of skills; they still fill less than 10 per cent of the places on information courses.

Improvements are already under way to increase the supply of key skills. However, an improved understanding of future manpower needs for all the new, and indeed existing, technologies is still needed to aid planners in government, education and industry, careers advisers and potential entrants. Such changing needs, and the different ways in which they might be met, such as retraining or substitution, also need to be regularly monitored and reviewed in a systematic way; and employment, education and training policies adjusted as appropriate if skill shortages are not to be a recurring problem in the future. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Brighton BN1 9RF, UK.